

STIC-Biotech/ChemLib

154881

Mej

From: Dunston, Jennifer
Sent: Tuesday, May 31, 2005 10:53 AM
To: STIC-Biotech/ChemLib
Subject: Sequence Search 10/826523

Please do a sequence search for the nucleic acid sequences of SEQ ID NO: 40 and SEQ ID NO: 41 against the commercial and interference nucleotide databases.

The length of SEQ ID NO: 40 is 707 nucleotides. The length of SEQ ID NO: 41 is 3662 nucleotides. The sequences are related in that SEQ ID NO: 41 comprises the sequence of SEQ ID NO: 40.

Thank you.

Jennifer Dunston, Ph.D.
USPTO Art Unit 1636
REM 2B76
Mailbox: REM 2C70
(571) 272-2916

RECEIVED
MAY 31 2005
STIC

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: 6-6
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 19:22:55 ; Search time 385.136 Seconds
(without alignments)
11284.911 Million cell updates/sec

Title: US-10-826-523-40

Perfect score: 707

Sequence: 1 ggaatccatcgctcaatttt.....aaattgacgcatggatccc 707

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5706582 seqs, 3073711274 residues

Total number of hits satisfying chosen parameters: 11413164

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*

20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*

21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	707	100.0	707	13	US-10-001-189-40
2	707	100.0	4613	13	US-10-001-189-46
3	707	100.0	8999	13	US-10-001-189-48
4	707	100.0	9012	13	US-10-001-189-49
5	707	100.0	9013	13	US-10-001-189-50
6	685	96.9	4943	13	US-10-001-189-54
7	685	96.9	4944	13	US-10-001-189-55
8	685	96.9	4944	13	US-10-001-189-56
9	566.2	80.1	26565	16	US-10-085-959-91
10	566.2	80.1	46819	14	US-10-114-170-72
11	561.4	79.4	22306	14	US-10-114-170-251
					Sequence 40, Appl
					Sequence 46, Appl
					Sequence 48, Appl
					Sequence 49, Appl
					Sequence 50, Appl
					Sequence 54, Appl
					Sequence 55, Appl
					Sequence 56, Appl
					Sequence 91, Appl
					Sequence 72, Appl
					Sequence 251, Appl

12	519.2	73.4	3662	13	US-10-001-189-41	Sequence 41, Appl
c 13	511	72.3	4941	13	US-10-001-189-53	Sequence 53, Appl
c 14	511	72.3	4951	13	US-10-001-189-51	Sequence 51, Appl
c 15	511	72.3	4952	13	US-10-001-189-52	Sequence 52, Appl
c 16	425.8	60.2	1427	18	US-10-363-345A-37517	Sequence 37517, A
c 17	425.8	60.2	1427	18	US-10-363-345A-37518	Sequence 37518, A
c 18	425.8	60.2	1427	18	US-10-363-483A-37517	Sequence 37517, A
c 19	425.8	60.2	1427	19	US-10-363-483A-37518	Sequence 37518, A
c 20	382.6	54.1	1427	18	US-10-363-345A-37519	Sequence 37519, A
c 21	382.6	54.1	1427	18	US-10-363-345A-37520	Sequence 37520, A
c 22	382.6	54.1	1427	19	US-10-363-483A-37519	Sequence 37519, A
c 23	382.6	54.1	1427	19	US-10-363-483A-37520	Sequence 37520, A
c 24	223	31.5	34063	14	US-10-114-170-96	Sequence 96, Appl
c 25	222.2	31.4	72480	17	US-10-418-837-2	Sequence 2, Appl
c 26	96	13.6	847	10	US-09-798-889-27	Sequence 27, Appl
c 27	96	13.6	847	17	US-10-633-680-27	Sequence 27, Appl
c 28	78	11.0	9984	13	US-10-001-189-63	Sequence 63, Appl
c 29	72.8	10.3	7670	13	US-10-001-189-57	Sequence 57, Appl
c 30	40	5.7	43	13	US-10-001-189-19	Sequence 19, Appl
c 31	40	5.7	3089	18	US-10-425-115-49726	Sequence 49726, A
c 32	39.8	5.6	676	18	US-10-425-115-91844	Sequence 91844, A
c 33	39.8	5.6	1673	17	US-10-425-114-21174	Sequence 21174, A
c 34	39.8	5.6	1796	17	US-10-425-114-29053	Sequence 29053, A
c 35	39.8	5.6	1872	18	US-10-425-115-91842	Sequence 91842, A
c 36	39	5.5	39	13	US-10-001-189-16	Sequence 16, Appl
c 37	38.4	5.4	1473	17	US-10-282-122A-20134	Sequence 20134, A
c 38	36.8	5.2	4187	18	US-10-437-963-91699	Sequence 91699, A
c 39	36.2	5.1	549	19	US-10-660-811A-23	Sequence 23, Appl
c 40	36.2	5.1	1403	18	US-10-767-701-13724	Sequence 13724, A
c 41	36.2	5.1	84830	19	US-10-660-811A-2	Sequence 2, Appl
c 42	36	5.1	1246	17	US-10-369-493-27114	Sequence 27114, A
c 43	35.4	5.0	2073	11	US-09-758-759-158	Sequence 158, App
c 44	35.4	5.0	96587	11	US-09-997-722-250	Sequence 250, App
c 45	35.4	5.0	109519	11	US-09-758-759-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-10-001-189-40

Sequence 40, Application US/10001189

Publication No. US20020173634A1

GENERAL INFORMATION:

APPLICANT: FRASER JR., MALCOLM J.

APPLICANT: LI, XU

APPLICANT: BEAM, TERESA

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION

TITLE OF INVENTION: VECTOR PIGGYBAC

FILE REFERENCE: 835910-92098

CURRENT APPLICATION NUMBER: US/10/001,189

CURRENT FILING DATE: 2001-10-30

PRIOR APPLICATION NUMBER: 60/244,984

PRIOR FILING DATE: 2000-11-01

PRIOR APPLICATION NUMBER: 60/244,677

PRIOR FILING DATE: 2000-10-31

NUMBER OF SEQ ID NOS: 70

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 40

LENGTH: 707

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: ITR Cartridge

OTHER INFORMATION: sequence

US-10-001-189-40

Query Match 100.0%; Score 707; DB 13; Length 707;
Best Local Similarity 100.0%; Pred. No. 1.8e-235;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCCATCGTCAATTTTACGACACTATCTTTCTAGGGTTAATCTAGCTCATCAGG 60

```
Db 1 G G A T C C C A T G C G T C A A T T T T A C G C A G A C T A T C T T C T A G G T T A A T C T A G C T C A T C A G G 60
Qy 61 A T C A T A T C G T C G G G T C T T T T T C C G G C T C A G T C A T C G C C A A G C C A G C G T A T C T G G G C A 120
Db 61 A T C A T A T C G T C G G G T C T T T T T C C G G C T C A G T C A T C G C C A A G C C A G C G T A T C T G G G C A 120
Qy 121 T C G G G A G A A G A A G C C G T G C C T T T T C C G C G A G G T T G A A G C G G C A T G A A A G A G T T T G 180
Db 121 T C G G G A G A A G A A G C C G T G C C T T T T C C G C G A G G T T G A A G C G G C A T G A A A G A G T T T G 180
Qy 181 C C G A G G A T G A C T G C T G C T G C A T T G A C G T T T G A G C G A A A A C G C A G T T T A C C A T G A T T C 240
Db 181 C C G A G G A T G A C T G C T G C T G C A T T G A C G T T T G A G C G A A A A C G C A G T T T A C C A T G A T T C 240
Qy 241 G G A A G G T G G C C A T G C A C G C C T T T A A C G G T A A C T G T T C G T T C A G G C C A C C T G G G A T A 300
Db 241 G G A A G G T G G C C A T G C A C G C C T T T A A C G G T A A C T G T T C G T T C A G G C C A C C T G G G A T A 300
Qy 301 C C A G T T C G T C G C G C T T T T C C G G A C A C A G T T C C G G A T G G T C A G C C G A A G C G C A T C A G C A 360
Db 301 C C A G T T C G T C G C G C T T T T C C G G A C A C A G T T C C G G A T G G T C A G C C G A A G C G C A T C A G C A 360
Qy 361 A C C G G A A C A A T A C C G G C G A C A G C C G G A A C T G C G T G C C G T G T G C A G A T T A A T G A C A G C G 420
Db 361 A C C G G A A C A A T A C C G G C G A C A G C C G G A A C T G C G T G T G C A G A T T A A T G A C A G C G 420
Qy 421 G T C G C G C G T G G A T A T T A C T C A G C G A G A C G G G T A T C T G C T G A T G C C C A G A A A T 480
Db 421 G T C G C G C G T G G A T A T T A C T C A G C G A G A C G G G T A T C T G C T G A T G C C C A G A A A T 480
Qy 481 G G A C A T G G A T A C C C G T G A G T T A C C C G G C G C C C T G T T C A T T C A G T T T T T G A A C 540
Db 481 G G A C A T G G A T A C C C G T G A G T T A C C C G G C G C C C T G T T C A T T C A G T T T T T G A A C 540
Qy 541 C C G T G A G G A C G G G C A G A C T C G C G T G C A A A T G T T T T A C A G C G T G A T G G A G C A G A T G A 600
Db 541 C C G T G A G G A C G G G C A G A C T C G C G T G C A A A T G T T T T A C A G C G T G A T G G A G C A G A T G A 600
Qy 601 A G A T G C T C A C A C G C T G C A A C A C A C G C A G C T A G A T T A A C C T A G A A G A T A A T C A T A T T G 660
Db 601 A G A T G C T C A C A C G C T G C A A C A C A C G C A G C T A G A T T A A C C T A G A A G A T A A T C A T A T T G 660
Qy 661 T G A C G T A C G T T A A G A T A A T C A T G C G T A A A A T T G A C G C A T G G G A T C C 707
Db 661 T G A C G T A C G T T A A G A T A A T C A T G C G T A A A A T T G A C G C A T G G G A T C C 707
```

RESULT 2

```
US-10-001-189-46
; Sequence 46, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 4613
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: Description of Artificial Sequence: pCRII-ITR
;
; OTHER INFORMATION: sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (344)..(922)
US-10-001-189-46
```

```
Query Match 100.0%; Score 707; DB 13; Length 4613;
Best Local Similarity 100.0%; Pred No. 4,9e-235;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 G G A T C C C A T G C G T C A A T T T T A C G C A G A C T A T C T T C T A G G T T A A T C T A G C T C A T C A G G 60
Db 294 G G A T C C C A T G C G T C A A T T T T A C G C A G A C T A T C T T C T A G G T T A A T C T A G C T C A T C A G G 353
Qy 61 A T C A T A T C G T C G G G T C T T T T T C C G G C T C A G T C A T C G C C A A G C C A G C T G C G C T A T C T G G C A 120
Db 354 A T C A T A T C G T C G G G T C T T T T T C C G G C T C A G T C A T C G C C A A G C C A G C T G C G C T A T C T G G C A 413
Qy 121 T C G G G A G A A G A A G C C G T G C C T T T T C C G C G A G G T T G A A G C G G C A T G A A A G A G T T T G 180
Db 414 T C G G G A G A A G A A G C C G T G C C T T T T C C G C G A G G T T G A A G C G G C A T G A A A G A G T T T G 473
Qy 181 C C G A G A T G A C T G C T G C T G C A T T G A C G T T T G A G C G A A A A C G C A C G T T T A C C A T G A T T C 240
Db 474 C C G A G A T G A C T G C T G C T G C A T T G A C G T T T G A G C G A A A A C G C A C G T T T A C C A T G A T T C 533
Qy 241 G G A A G G T G T G C C A T G C A C G C C T T T A A C G G T G A A C T G T T C G T T C A G G C C A C C T G G G A T A 300
Db 534 G G A A G G T G T G C C A T G C A C G C C T T T A A C G G T G A A C T G T T C G T T C A G G C C A C C T G G G A T A 593
Qy 301 C C A G T T C G T C G C G C T T T T C C G G A C A C A G T T C C G A T G T C A G C C C G A A G C G C A T C A G C A 360
Db 594 C C A G T T C G T C G C G C T T T T C C G G A C A C A G T T C C G A T G T C A G C C C G A A G C G C A T C A G C A 653
Qy 361 A C C G G A A C A A T A C C G G C G A C A G C C G G A A C T G C G T G C C G T G C A G A T T A A T G A C A G C G 420
Db 654 A C C G G A A C A A T A C C G G C G A C A G C C G G A A C T G C G T G C C G T G T G C A G A T T A A T G A C A G C G 713
Qy 421 G T C G C G C G T G G A T A T T A C G T C A G C G A G A C G G G T A T C T G C T G A T G C C C A G A A A T 480
Db 714 G T C G C G C G T G G A T A T T A C G T C A G C G A G A C G G G T A T C T G C T G A T G C C C A G A A A T 773
Qy 481 G G A C A T G G A T A C C C G T G A G T T A C C C G G C G C C C T G T T C A T T C A G T T T T T G A A C 540
Db 774 G G A C A T G G A T A C C C G T G A G T T A C C C G G C G C C C T G T T C A T T C A G T T T T T G A A C 833
Qy 541 C C G T G A G A C C G G G C A G A C T C G C G T G C A A A T G T G T T T T A C A G C G T G A T G G A G C A G A T G A 600
Db 834 C C G T G A G A C C G G G C A G A C T C G C G T G C A A A T G T G T T T T A C A G C G T G A T G G A G C A G A T G A 893
Qy 601 A G A T G C T C A C A C G C T G C A A C A C A C G C A G C T A G A T T A A C C T A G A A G A T A A T C A T A T T G 660
Db 894 A G A T G C T C A C A C G C T G C A A C A C A C G C A G C T A G A T T A A C C T A G A A G A T A A T C A T A T T G 953
Qy 661 T G A C G T A C G T T A A G A T A A T C A T G C G T A A A A T T G A C G C A T G G G A T C C 707
Db 954 T G A C G T A C G T T A A G A T A A T C A T G C G T A A A A T T G A C G C A T G G G A T C C 1000
```

RESULT 3

```
US-10-001-189-48/c
; Sequence 48, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
```


; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR FILING DATE: 2000-11-01
; PRIOR FILING DATE: 2000-11-01
; PRIOR FILING DATE: 2000-11-01
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 8999
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-EYFP
; OTHER INFORMATION: sequence
US-10-001-189-48

Query Match 100.0%; Score 707; DB 13; Length 8999;
Best Local Similarity 100.0%; Pred. No. 6.9e-235;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCCATCGCTCAATTTTACGACAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
DB 8950 GGATCCCATCGCTCAATTTTACGACAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 8891

QY 61 ATCATATCGTGGGTCTTTTTCGGCTCAGTATCGCCCAAGCTGGCGTATCTGGGCA 120
DB 8890 ATCATATCGTGGGTCTTTTTCGGCTCAGTATCGCCCAAGCTGGCGTATCTGGGCA 8831

QY 121 TCGGGGAGGAAGAACGCGCTTTTTCGGCGAGGTTGAAGCGGCAATGAAGAGTTTG 180
DB 8830 TCGGGGAGGAAGAACGCGCTTTTTCGGCGAGGTTGAAGCGGCAATGAAGAGTTTG 8771

QY 181 CCAGGATGATCTGCTGCTGATTCAGCTTGAGCGAAACGACGTTTACCATGATTC 240
DB 8770 CCAGGATGATCTGCTGCTGATTCAGCTTGAGCGAAACGACGTTTACCATGATTC 8711

QY 241 GGAAGGTGGGCGATGACGCTTTTAAACGGTGAATCTGTTTCAGGCCACCTGGGATA 300
DB 8710 GGAAGGTGGGCGATGACGCTTTTAAACGGTGAATCTGTTTCAGGCCACCTGGGATA 8651

QY 301 CCAGTTCGTCGGGCTTTTTCGGGACACAGTTCGGATGGTTCAGCCGAGCGATCAGCA 360
DB 8650 CCAGTTCGTCGGGCTTTTTCGGGACACAGTTCGGATGGTTCAGCCGAGCGATCAGCA 8591

QY 361 ACCCGAACATACCGGCGACAGCGGAACTCCGCTGCGGTGTCAGATTAATGACAGCG 420
DB 8590 ACCCGAACATACCGGCGACAGCGGAACTCCGCTGCGGTGTCAGATTAATGACAGCG 8531

QY 421 GTGCGGCTGGGATATTACGTCAGCGAGGACGGGTATCTGGCTGGATCCCGAGAAAT 480
DB 8530 GTGCGGCTGGGATATTACGTCAGCGAGGACGGGTATCTGGCTGGATCCCGAGAAAT 8471

QY 481 GGACATGGATACCCCGTGGTATCCCGGCGGCGGCTCGTTCATTCAGCTTTTGAAC 540
DB 8470 GGACATGGATACCCCGTGGTATCCCGGCGGCGGCTCGTTCATTCAGCTTTTGAAC 8411

QY 541 CCCTGGAGGACGGGCGAGCTCGGGGTGCAAAATGTTTACAGCGTGTGATGGAGCAGATGA 600
DB 8410 CCCTGGAGGACGGGCGAGCTCGGGGTGCAAAATGTTTACAGCGTGTGATGGAGCAGATGA 8351

QY 601 AGATGTCGACGCTGCGGAGAACACGCGAGCTAGATTAACCTAGAAAGATTAATCATTTG 660
DB 8350 AGATGTCGACGCTGCGGAGAACACGCGAGCTAGATTAACCTAGAAAGATTAATCATTTG 8291

QY 661 TGACGTACGTTAAAGATAATCATCGTAAATTTGACGATGGATCC 707
DB 8290 TGACGTACGTTAAAGATAATCATCGTAAATTTGACGATGGATCC 8244

RESULT 4

US-10-001-189-49/c

; Sequence 49, Application US/10001189

; Publication No. US20020173634A1

; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR FILING DATE: 2000-11-01
; PRIOR FILING DATE: 2000-11-01
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 9012
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-ECFP
; OTHER INFORMATION: sequence
US-10-001-189-49

Query Match 100.0%; Score 707; DB 13; Length 9012;
Best Local Similarity 100.0%; Pred. No. 6.9e-235;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCCATCGCTCAATTTTACGACAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
DB 8963 GGATCCCATCGCTCAATTTTACGACAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 8904

QY 61 ATCATATCGTGGGTCTTTTTCGGCTCAGTATCGCCCAAGCTGGCGTATCTGGGCA 120
DB 8903 ATCATATCGTGGGTCTTTTTCGGCTCAGTATCGCCCAAGCTGGCGTATCTGGGCA 8844

QY 121 TCGGGGAGGAAGAACGCGCTTTTTCGGCGAGGTTGAAGCGGCAATGAAGAGTTTG 180
DB 8843 TCGGGGAGGAAGAACGCGCTTTTTCGGCGAGGTTGAAGCGGCAATGAAGAGTTTG 8784

QY 181 CCAGGATGATCTGCTGCTGATTCAGTTCAGCGAAACGACGTTTACCATGATTC 240
DB 8783 CCAGGATGATCTGCTGCTGATTCAGTTCAGCGAAACGACGTTTACCATGATTC 8724

QY 241 GGAAGGTGGGCGATGACGCTTTTAAACGGTGAATCTGTTTCAGGCCACCTGGGATA 300
DB 8723 GGAAGGTGGGCGATGACGCTTTTAAACGGTGAATCTGTTTCAGGCCACCTGGGATA 8664

QY 301 CCAGTTCGTCGGGCTTTTTCGGGACACAGTTCGGATGGTTCAGCCGAGCGATCAGCA 360
DB 8663 CCAGTTCGTCGGGCTTTTTCGGGACACAGTTCGGATGGTTCAGCCGAGCGATCAGCA 8604

QY 361 ACCCGAACATACCGGCGACAGCGGAACTCCGCTGCGGTGTCAGATTAATGACAGCG 420
DB 8603 ACCCGAACATACCGGCGACAGCGGAACTCCGCTGCGGTGTCAGATTAATGACAGCG 8544

QY 421 GTGCGGCTGGGATATTACGTCAGCGAGGACGGGTATCTGGCTGGATCCCGAGAAAT 480
DB 8543 GTGCGGCTGGGATATTACGTCAGCGAGGACGGGTATCTGGCTGGATCCCGAGAAAT 8484

QY 481 GGACATGGATACCCCGTGGTATCCCGGCGGCGGCTCGTTCATTCAGCTTTTGAAC 540
DB 8483 GGACATGGATACCCCGTGGTATCCCGGCGGCGGCTCGTTCATTCAGCTTTTGAAC 8424

QY 541 CCCTGGAGGACGGGCGAGCTCGGGGTGCAAAATGTTTTCAGCGTGTGATGGAGCAGATGA 600
DB 8423 CCCTGGAGGACGGGCGAGCTCGGGGTGCAAAATGTTTTCAGCGTGTGATGGAGCAGATGA 8364

QY 601 AGATGTCGACGCTGCGGAGAACACGCGAGCTAGATTAACCTAGAAAGATTAATCATTTG 660
DB 8363 AGATGTCGACGCTGCGGAGAACACGCGAGCTAGATTAACCTAGAAAGATTAATCATTTG 8304

Qy	661	TGACGTACGTTAAAGATAATCATCGTAAAAATTGACGCATGGGATCC	707
Dd	8303	TGACGTACGTTAAAGATAATCATCGTAAAAATTGACGCATGGGATCC	8257
RESULT 5			
US-10-001-189-50/c			
; Sequence 50, Application US/10001189			
; Publication No. US20020173634A1			
; GENERAL INFORMATION:			
; APPLICANT: FRASER JR., MALCOLM J.			
; APPLICANT: LI, XU			
; APPLICANT: BEAM, TERESA			
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING			
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION			
; FILE REFERENCE: 835910-92098			
; CURRENT APPLICATION NUMBER: US/10/001,189			
; PRIOR FILING DATE: 2001-10-30			
; PRIOR APPLICATION NUMBER: 60/244,984			
; PRIOR FILING DATE: 2000-11-01			
; PRIOR APPLICATION NUMBER: 60/244,677			
; PRIOR FILING DATE: 2000-10-31			
; NUMBER OF SEQ ID NOS: 70			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 50			
; LENGTH: 9013			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-EGFP			
; OTHER INFORMATION: sequence			
US-10-001-189-50			
Query Match 100.0%; Score 707; DB 13; Length 9013;			
Best Local Similarity 100.0%; Pred. No. 6.9e-235;			
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	GGATCCCATGCGTCGAATTTTAACGACACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG	60
Dd	8964	GGATCCCATGCGTCGAATTTTAACGACACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG	8905
Qy	61	ATCATATCGTCGGGTCCTTTTTTCGGCTCAGTCATCGGCCAACGCTGGCGCTATCTGGGCA	120
Dd	8904	ATCATATCGTCGGGTCCTTTTTTCGGCTCAGTCATCGGCCAACGCTGGCGCTATCTGGGCA	8845
Qy	121	TCGGGAGGAGAAGACCCTGCTCTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG	180
Dd	8844	TCGGGAGGAGAAGACCCTGCTCTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG	8785
Qy	181	CCGAGGATGACTGCTGCTGCATTGAGCGAAAACGCAACGCTTTACCATGATGATTC	240
Dd	8784	CCGAGGATGACTGCTGCTGCATTGAGCGAAAACGCAACGCTTTACCATGATGATTC	8725
Qy	241	GGAAGGTTGGCCATGACCGCTTTTAACGGTGAACCTGTTCCAGCGCACCTGGGATA	300
Dd	8724	GGAAGGTTGGCCATGACCGCTTTTAACGGTGAACCTGTTCCAGCGCACCTGGGATA	8665
Qy	301	CCAGTTTCGTCCGGCTTTTCCGACACAGTTCGGGATGTCAGCCCCGAAGCGCATCAGCA	360
Dd	8664	CCAGTTTCGTCCGGCTTTTCCGACACAGTTCGGGATGTCAGCCCCGAAGCGCATCAGCA	8605
Qy	361	ACCCGAACAATACCGCGGACAGCCGGAATCGCCGTCCGGTGTGCAGATTAAATGACACGC	420
Dd	8604	ACCCGAACAATACCGCGGACAGCCGGAATCGCCGTCCGGTGTGCAGATTAAATGACACGC	8545
Qy	421	GTCCGGCGCTGGGATATTACGTCCAGGAGAGCGGTATCTCTGGCTGGATGCGCAGAAAT	480
Dd	8544	GTCCGGCGCTGGGATATTACGTCCAGGAGAGCGGTATCTCTGGCTGGATGCGCAGAAAT	8485
Qy	481	GGAATGGATACCCCCTGAGTTTACCCGGCGGCGCCCTCGTTCAATCACTGTTTGAAC	540
Dd	8484	GGAATGGATACCCCCTGAGTTTACCCGGCGGCGCCCTCGTTCAATCACTGTTTGAAC	8425

Qy	541	CCGTGGAGGACGGGCGAGACTCGCGTGCAAAATGTGTTTACAGCGTGATGGAGCAGATGA	600
Dd	8424	CCGTGGAGGACGGGCGAGACTCGCGTGCAAAATGTGTTTACAGCGTGATGGAGCAGATGA	8365
Qy	601	AGATGCTCGACACGCTGCGAGAACACGACGCTAGATTTAACCCCTAGAAAGATAATCATATTG	660
Dd	8364	AGATGCTCGACACGCTGCGAGAACACGACGCTAGATTTAACCCCTAGAAAGATAATCATATTG	8305
Qy	661	TGACGTACGTTAAAGATAATCATCGTAAAAATTGACGCATGGGATCC	707
Dd	8304	TGACGTACGTTAAAGATAATCATCGTAAAAATTGACGCATGGGATCC	8258
RESULT 6			
US-10-001-189-54			
; Sequence 54, Application US/10001189			
; Publication No. US20020173634A1			
; GENERAL INFORMATION:			
; APPLICANT: FRASER JR., MALCOLM J.			
; APPLICANT: LI, XU			
; APPLICANT: BEAM, TERESA			
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING			
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION			
; FILE REFERENCE: 835910-92098			
; CURRENT APPLICATION NUMBER: US/10/001,189			
; CURRENT FILING DATE: 2001-10-30			
; PRIOR APPLICATION NUMBER: 60/244,984			
; PRIOR FILING DATE: 2000-11-01			
; PRIOR APPLICATION NUMBER: 60/244,677			
; PRIOR FILING DATE: 2000-10-31			
; NUMBER OF SEQ ID NOS: 70			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 54			
; LENGTH: 4943			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: Description of Artificial Sequence: PBS-ITR-ECFP			
; OTHER INFORMATION: sequence			
US-10-001-189-54			
Query Match 96.9%; Score 685; DB 13; Length 4943;			
Best Local Similarity 99.2%; Pred. No. 2.4e-227;			
Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;			
Qy	1	GGATCCCATGCGTCGAATTTTAACGACACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG	60
Dd	719	GGATCCCATGCGTCGAATTTTAACGACACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG	778
Qy	61	ATCATATCGTCGGGTCCTTTTTTCGGGCTCAGTCATCGCCCCAAGCTGGCGCTATCTGGGCA	120
Dd	779	ATCATATCGTCGGGTCCTTTTTTCGGGCTCAGTCATCGCCCCAAGCTGGCGCTATCTGGGCA	838
Qy	121	TCGGGAGGAGAAGACCCTGCTCTTTCCCGGAGGTTGAAGCGGCATGGAAGAGTTTG	180
Dd	839	TCGGGAGGAGAAGACCCTGCTCTTTCCCGGAGGTTGAAGCGGCATGGAAGAGTTTG	898
Qy	181	CCGAGGATGACTGCTGCTGCATTGAGCGAAAACGCAACGCTTTACCATGATGATTC	240
Dd	899	CCGAGGATGACTGCTGCTGCATTGAGCGAAAACGCAACGCTTTACCATGATGATTC	958
Qy	241	GGAAGGTTGGCCATGACCGCTTTTAACGGTGAACCTGTTCCAGCGCACCTGGGATA	300
Dd	959	GGAAGGTTGGCCATGACCGCTTTTAACGGTGAACCTGTTCCAGCGCACCTGGGATA	1018
Qy	301	CCAGTTTCGTCCGGCTTTTCCGACACAGTTCGGGATGTCAGCCCCGAAGCGCATCAGCA	360
Dd	1019	CCAGTTTCGTCCGGCTTTTCCGACACAGTTCGGGATGTCAGCCCCGAAGCGCATCAGCA	1078
Qy	361	ACCCGAACAATACCGCGGACAGCCGGAATCGCCGTCCGGTGTGCAGATTAAATGACACGC	420

Db 1079 ACCGGAACAATACCGGCGACAGCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 1138
Qy 421 GTGCGGCGCTGGGATATTACGTTCAGCGAGGACGGGTATCTCGCTGGATCCGCGAGAAAT 480
Db 1139 GTGCGGCGCTGGGATATTACGTTCAGCGAGGACGGGTATCTCGCTGGATCCGCGAGAAAT 1198
Qy 481 GGACATGATACCGCTGAGTTACCGGCGGCGCGCTCGTTCAATTCAGCTTTTGAAC 540
Db 1199 GGACATGATACCGCTGAGTTACCGGCGG-----CTCGTTCAATTCAGCTTTTGAAC 1252
Qy 541 CCGTGAGGACGGGACAGCTCGCGGTCAAAATGTGTTTTACAGCGTGTGATGAGCAGATGA 600
Db 1253 CCGTGAGGACGGGACAGCTCGCGGTCAAAATGTGTTTTACAGCGTGTGATGAGCAGATGA 1312
Qy 601 AGATGCTCGACACGCTGCGAGACAGCAGCTAGATTAAACCTAGAAAGATAATCATATTG 660
Db 1313 AGATGCTCGACACGCTGCGAGACAGCAGCTAGATTAAACCTAGAAAGATAATCATATTG 1372
Qy 661 TGACGTACGTTAAAGATAATCATGCGTAAATTTGACGCGATGGATCC 707
Db 1373 TGACGTACGTTAAAGATAATCATGCGTAAATTTGACGCGATGGATCC 1419

RESULT 7

US-10-001-189-55
; Sequence 55, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PBS-ITR-EGFP
; OTHER INFORMATION: sequence
US-10-001-189-55

Query Match 96.9%; Score 685; DB 13; Length 4944;
Best Local Similarity 99.2%; Pred. No. 2.4e-227;
Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;
Qy 1 GGATCCCATCGCTCAATTTTACGCGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db 719 GGATCCCATCGCTCAATTTTACGCGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 778
Qy 61 ATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATGCCCAAGCTGGCGCTATCTGGGCA 120
Db 779 ATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATGCCCAAGCTGGCGCTATCTGGGCA 838
Qy 121 TCGGGGAGGAGAACGCCGTCCTTTTCCGCGAGGTTGAAGCGGCATGGAAGAGTTTG 180
Db 839 TCGGGGAGGAGAACGCCGTCCTTTTCCGCGAGGTTGAAGCGGCATGGAAGAGTTTG 898
Qy 181 CCAGGATGATCTGCTGCTGATTGACGTTGAGCGGAAACGACGCTTTTACCATCATGATTC 240
Db 899 CCAGGATGATCTGCTGCTGATTGACGTTGAGCGGAAACGACGCTTTTACCATCATGATTC 958
Qy 241 GGGAAAGGTGTGGCCATGACACGCTTTTAAACGGTGAACCTGTTCTGTCAGGCCACCTGGGATA 300

Db 959 GGGAAAGGTGTGGCCATGACGCTTTTAAACGGTGAACCTGTTCTGTCAGGCCACCTGGGATA 1018
Qy 301 CCAAGTTTCGTCCGGCTTTTCCGAGACAGTTTCCGGATGGTTCAGCCCAAGCGCATCAGCA 360
Db 1019 CCAAGTTTCGTCCGGCTTTTCCGAGACAGTTTCCGGATGGTTCAGCCCAAGCGCATCAGCA 1078
Qy 361 ACCCGAAACAATACCGGCGACAGCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 420
Db 1079 ACCCGAAACAATACCGGCGACAGCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 1138
Qy 421 GTGCGGCGCTGGGATATTACGTTCAGCGAGGACGGGTATCTCGCTGGATCCGCGAGAAAT 480
Db 1139 GTGCGGCGCTGGGATATTACGTTCAGCGAGGACGGGTATCTCGCTGGATCCGCGAGAAAT 1198
Qy 481 GGACATGATACCGCTGAGTTACCGGCGGCGCGCTCGTTCAATTCAGCTTTTGAAC 540
Db 1199 GGACATGATACCGCTGAGTTACCGGCGG-----CTCGTTCAATTCAGCTTTTGAAC 1252
Qy 541 CCGTGAGGACGGGACAGCTCGCGGTGCAAAATGTGTTTTACAGCGTGTGATGAGCAGATGA 600
Db 1253 CCGTGAGGACGGGACAGCTCGCGGTGCAAAATGTGTTTTACAGCGTGTGATGAGCAGATGA 1312
Qy 601 AGATGCTCGACACGCTGCGAGACAGCAGCTAGATTAAACCTAGAAAGATAATCATATTG 660
Db 1313 AGATGCTCGACACGCTGCGAGACAGCAGCTAGATTAAACCTAGAAAGATAATCATATTG 1372
Qy 661 TGACGTACGTTAAAGATAATCATGCGTAAATTTGACGCGATGGATCC 707
Db 1373 TGACGTACGTTAAAGATAATCATGCGTAAATTTGACGCGATGGATCC 1419

RESULT 8

US-10-001-189-56
; Sequence 56, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 4944
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PBS-ITR-EYFP
; OTHER INFORMATION: sequence
US-10-001-189-56

Query Match 96.9%; Score 685; DB 13; Length 4944;
Best Local Similarity 99.2%; Pred. No. 2.4e-227;
Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;
Qy 1 GGATCCCATCGCTCAATTTTACGCGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db 719 GGATCCCATCGCTCAATTTTACGCGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 778
Qy 61 ATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATGCCCAAGCTGGCGCTATCTGGGCA 120
Db 779 ATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATGCCCAAGCTGGCGCTATCTGGGCA 838


```
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 72:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46819
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-10-114-170-72

Query Match      80.1%; Score 566.2; DB 14; Length 46819;
Best Local Similarity 97.8%; Pred. No. 2.3e-185;
Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCAATCGCCCAAGC 104
Db 14795 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCAATCGCCCAAGC 14854

QY 105 TGGCGCTATCTGGGATCGGGAGGAAGACCGCGTCCCTTTTCCCGGAGGTGAAGCG 164
Db 14855 TGGCGCTATCTGGGATCGGGAGGAAGACCGCGTCCCTTTTCCCGGAGGTGAAGCG 14914

QY 165 GCATGAAAGAGTTTGGCGAGGATGACTGCTGCTGCAATGACGTTGAGCGAAGAACGACG 224
Db 14915 GCATGAAAGAGTTTGGCGAGGATGACTGCTGCTGCAATGACGTTGAGCGAAGAACGACG 14974

QY 225 TTTACCATGATGATTCGGGAAGGTGGCCATGCAAGCCCTTTTAAACCGTGAACCTGTTGCTT 284
Db 14975 TTTACCATGATGATTCGGGAAGGTGGCCATGCAAGCCCTTTTAAACCGTGAACCTGTTGCTT 15034

QY 285 CAGGCCACCTGGGATACCACTGCTCGCGCTTTTCGGGACACAGTTCGGGATGGTCAGC 344
Db 15035 CAGGCCACCTGGGATACCACTGCTCGCGCTTTTCGGGACACAGTTCGGGATGGTCAGC 15094

QY 345 CCGAAGGCGATCAGCAACCGAACAATACCGGGACAGCGGAACTCCCGTCCGGGTG 404
Db 15095 CCGAAGGCGATCAGCAACCGAACAATACCGGGACAGCGGAACTCCCGTCCGGGTG 15154

QY 405 CAGATTAAATGACAGCGGTGGCGCTGGGATATTTACGTCAGCGAGGACGGGTATCTCTGGC 464
Db 15155 CAGATTAAATGACAGCGGTGGCGCTGGGATATTTACGTCAGCGAGGACGGGTATCTCTGGC 15214

QY 465 TGGATGCGCAGAAATGGAATGATGATACCCCGTGAAGTTACCGCGGCGCGCTGCTTTC 524
Db 15215 TGGATGCGCAGAAATGGAATGATGATACCCCGTGAAGTTACCGCGGCGCGCTGCTTTC 15274

QY 525 ATTACGCTTTTGAACCCGTTGGAGGAGCGGACGACTCGCGGTGCAATGTTTACAGC 584
Db 15275 ATTACGCTTTTGAACCCGTTGGAGGAGCGGACGACTCGCGGTGCAATGTTTACAGC 15334

QY 585 GTGATGAGCAGATGAAGATGCTCGACACGCTGCGAAGAACGACGCT 631
Db 15335 GTGATGAGCAGATGAAGATGCTCGACACGCTGCGAAGAACGACGCT 15381
```

RESULT 11

```
US-10-114-170-251/c
; Sequence 251, Application US/10114170
; Publication No. US20030023075A1
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
;            Burland, Valerie
;            Perna, Nicole T.
;            Plunkett, Guy
;            Welch, Rod
; TITLE OF INVENTION: No. US20030023075A1el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/114,170
; FILING DATE: 01-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/453,702
; FILING DATE: 03-DEC-1999
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 251:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22306
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 251:
US-10-114-170-251

Query Match      79.4%; Score 561.4; DB 14; Length 22306;
Best Local Similarity 97.3%; Pred. No. 7.3e-184;
Matches 571; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCAATCGCCCAAGC 104
Db 17193 ATCCAGTTCGATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCAATCGCCCAAGC 17134

QY 105 TGGCGCTATCTGGGATCGGGAGGAAGACCGCGTCCCTTTTCCCGGAGGTGAAGCG 164
Db 17133 TGGCGCTATCTGGGATCGGGAGGAAGACCGCGTCCCTTTTCCCGGAGGTGAAGCG 17074

QY 165 GCATGAAAGAGTTTGGCGAGGATGACTGCTGCTGCAATGACGTTGAGCGAAGAACGACG 224
Db 17073 GCATGAAAGAGTTTGGCGAGGATGACTGCTGCTGCAATGACGTTGAGCGAAGAACGACG 17014

QY 225 TTTACCATGATGATTCGGGAAGGTGGCCATGCAAGCCCTTTTAAACCGTGAACCTGTTGCTT 284
Db 17013 TTTACCATGATGATTCGGGAAGGTGGCCATGCAAGCCCTTTTAAACCGTGAACCTGTTGCTT 16954

QY 285 CAGGCCACCTGGGATACCACTGCTCGCGCTTTTCCGGACACAGTTCGGGATGGTCAGC 344
Db 16953 CAGGCCACCTGGGATACCACTGCTCGCGCTTTTCCGGACACAGTTCGGGATGGTCAGC 16894

QY 345 CCGAAGGCGATCAGCAACCGAACAATACCGGGACAGCGGAACTCCCGTCCGGGTG 404
Db 16893 CCGAAGGCGATCAGCAACCGAACAATACCGGGACAGCGGAACTCCCGTCCGGGTG 16834

QY 405 CAGATTAAATGACAGCGGTGGCGCTGGGATATTTACGTCAGCGAGGACGGGTATCTCTGGC 464
Db 16833 CAGATTAAATGACAGCGGTGGCGCTGGGATATTTACGTCAGCGAGGACGGGTATCTCTGGC 16774

QY 465 TGGATGCGCAGAAATGGAATGATGATACCCCGTGAAGTTACCCGCGCGCGCTGCTTTC 524
Db 16773 TGGATGCGCAGAAATGGAATGATGATACCCCGTGAAGTTACCCGCGCGCGCTGCTTTC 16714
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 13:34:58 ; Search time 342.253 Seconds
(without alignments)
12228.534 Million cell updates/sec

Title: US-10-826-523-40
Perfect score: 707
Sequence: 1 ggatccatcggtcaatttt.....aaattgacgatgggatcc 707

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002as.*
7: Geneseq2002bs.*
8: Geneseq2003as.*
9: Geneseq2003bs.*
10: Geneseq2003cs.*
11: Geneseq2003ds.*
12: Geneseq2004as.*
13: Geneseq2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	707	100.0	707	10	ACA55353 Eukaryoti
2	707	100.0	4613	10	ACA55360 Transform
C 3	705.4	99.8	8999	10	ACA55361 Transform
C 4	705.4	99.8	9009	10	ACA55362 Transform
C 5	705.4	99.8	9012	10	ACA55363 Transform
6	685	96.9	4943	10	ACA55369 Transform
7	685	96.9	4943	10	ACA55369 Transform
8	681.8	96.4	4944	10	ACA55367 Transform
9	585.4	82.8	48502	12	ACA55368 Transform
C 10	580.6	82.1	4164	4	Adp70043 Nucleic a
C 11	566.2	80.1	26565	6	Adh80491 Escherich
C 12	566.2	80.1	26565	10	Adh80491 Escherich
C 13	566.2	80.1	46819	9	Adc19059 E. coli 0
14	564.6	79.9	46897	10	Adc00585 Enterohae
C 15	561.4	79.4	22306	9	Adc19238 E. coli 0
C 16	561.4	79.4	91740	10	Adc00956 Enterohae
17	519.2	73.4	3661	10	ACA55354 Transform
C 18	511	72.3	4941	10	ACA55366 Transform
C 19	511	72.3	4952	10	ACA55365 Transform
C 20	499	70.6	4951	10	ACA55364 Transform

21	425.8	60.2	1427	6	ABQ50926
C 22	425.8	60.2	1427	6	ABQ50927
23	392.6	55.5	3822	5	AAS94485
C 24	382.6	54.1	1427	6	ABQ50928
25	382.6	54.1	1427	6	ABQ50929
C 26	223	31.5	34063	9	ACD19083
C 27	223	31.5	44029	10	ADC00710
28	223	31.5	49650	10	ADC00365
C 29	222.2	31.4	134141	6	ABN83487
C 30	171.2	24.2	822	5	AAS80174
C 31	96	13.6	847	2	AAS80174
C 32	78	11.0	9980	10	ACA55352
C 33	73.8	10.4	5194	4	AAC85595
C 34	73.8	10.4	5194	9	ADA09845
C 35	73.4	10.4	5194	4	AAC85596
C 36	73.4	10.4	5194	9	ADA09846
C 37	72.8	10.3	7670	10	ACA55351
C 38	67.8	9.6	2480	8	ABV76195
39	67.8	9.6	5679	4	AAC85602
40	67.8	9.6	5679	9	ADA09857
C 41	67.8	9.6	6723	4	AAC85593
42	67.8	9.6	6723	4	AAC85592
C 43	67.8	9.6	6723	9	ADA09843
44	67.8	9.6	6723	9	ADA09842
C 45	67.8	9.6	7560	4	AAC85599

ALIGNMENTS

RESULT 1
ACA55353
ID ACA55353 standard; DNA; 707 BP.
XX
AC ACA55353;
XX
DT 09-JUN-2003 (first entry)
XX
DE Eukaryotic transformation vector piggyBAC ITR cartridge.
XX
KW PiggyBac; transposon; eukaryotic transformation vector; ds;
KW transformed cell; transformed embryo; transgenic; ITR cartridge.
XX
OS Synthetic.
XX
PN US2002173634-A1.
XX
PD 21-NOV-2002.
XX
PF 30-OCT-2001; 2001US-00001189.
XX
PR 31-OCT-2000; 2000US-0244677P.
PR 01-NOV-2000; 2000US-0244984P.
XX
PA (PRAS/) FRASER M J.
PA (LIXX/) LI X.
PA (BEAM/) BEAM T.
XX (HUAV/) HUA-VAN A.
XX
PI Fraser MJ, Li X, Beam T, Hua-Van A;
XX WPI; 2003-352597/33;
XX
PT New DNA molecule in the transposon piggyBac, useful for transferring
PT genes into host cells or embryos for transforming the cells of embryos
PT that can be used in making transgenic organisms.
XX
PS Example 3; Fig 3(C1); 151pp; English.
XX
CC The invention describes a DNA molecule comprising at least 163
CC consecutive nucleotide base pairs of the 3' terminal region beginning at
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of

CC the piggyBac molecule. The region extends from the restriction site SacI
CC to the end of the piggyBac molecule. The DNA molecule in the transposon
CC piggyBac is useful for transferring genes into host cells or embryos for
CC transforming the cells of embryos. The transposon contains genes for
CC useful for developing or making transgenic organisms. This sequence
CC represents a minimal sequence cartridge of the eukaryotic transformation
CC vector piggyBac
XX
SQ Sequence 707 BP; 164 A; 174 C; 208 G; 161 T; 0 U; 0 Other;

Query Match 100.0%; Score 707; DB 10; Length 707;
Best Local Similarity 100.0%; Pred. No. 2.6e-213;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATCCCATCGCTCAATTTTACGACGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
DB 1 GGATCCCATCGCTCAATTTTACGACGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
QY 61 ATCATATCGTGGGTCTTTTTCGGGCTCAGTCAATCGCCCAAGCTGGCGTATCTGGGCA 120
DB 61 ATCATATCGTGGGTCTTTTTCGGGCTCAGTCAATCGCCCAAGCTGGCGTATCTGGGCA 120
QY 121 TCGGGGAGGAAGAGCCGCTCTTTTCCCGGCTCAGTCAATCGCCCAAGCTGGCGTATCTGGGCA 180
DB 121 TCGGGGAGGAAGAGCCGCTCTTTTCCCGGCTCAGTCAATCGCCCAAGCTGGCGTATCTGGGCA 180
QY 181 CCGAGGATGACTGCTGCTCAATTTTACGACGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 240
DB 181 CCGAGGATGACTGCTGCTCAATTTTACGACGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 240
QY 241 GGGAGGATGCTGGGCTCAATTTTACGACGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 300
DB 241 GGGAGGATGCTGGGCTCAATTTTACGACGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 300
QY 301 CCAGTTGCTCGGCTCTTTTCCGACACAGTTCGGATGTCGAGTCCGAGTTCGAGTCCGAGTTCGAG 360
DB 301 CCAGTTGCTCGGCTCTTTTCCGACACAGTTCGGATGTCGAGTCCGAGTTCGAGTCCGAGTTCGAG 360
QY 361 ACCGACATATACCGGACAGCCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 420
DB 361 ACCGACATATACCGGACAGCCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 420
QY 421 GTGGGCTGCTGGGATATTCGCTCAGCGAGGACGGTATCTTCTAGGGTTAATCTAGCTGCATCAGG 480
DB 421 GTGGGCTGCTGGGATATTCGCTCAGCGAGGACGGTATCTTCTAGGGTTAATCTAGCTGCATCAGG 480
QY 481 GGACATGATACCCCGTGAATTTACCGGCGGCGGCTCGTTCATTCAGGTTTTCGAGTTCGAGTTCGAG 540
DB 481 GGACATGATACCCCGTGAATTTACCGGCGGCGGCTCGTTCATTCAGGTTTTCGAGTTCGAGTTCGAG 540
QY 541 CCGTGGAGGACGGGACGCTCGCGGTGCAATGTGTTTACGCGTGAATGTTTACGCGTGAATGTTTACG 600
DB 541 CCGTGGAGGACGGGACGCTCGCGGTGCAATGTGTTTACGCGTGAATGTTTACGCGTGAATGTTTACG 600
QY 601 AGATGCTCGACAGCTCGACACACGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 660
DB 601 AGATGCTCGACAGCTCGACACACGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 660
QY 661 TGACGTACGTTAAAGATAATCATCGTGAATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 707
DB 661 TGACGTACGTTAAAGATAATCATCGTGAATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 707

RESULT 2
ACA55360
ID ACA55360 standard; DNA; 4613 BP.
XX
AC ACA55360;
XX
DT 06-JUN-2003 (first entry)
XX
DE Transformation vector piggyBAC related plasmid pCRII-ITR.
XX

KW PiggyBac; transposon; eukaryotic transformation vector; ds;
KW transformed cell; transgenic embryo; transgenic; plasmid; cyclic;
KW circular.
XX
OS Synthetic.
XX
PN US2002173634-A1.
XX
PD 21-NOV-2002.
XX
XX 30-OCT-2001; 2001US-00001189.
PF
XX 31-OCT-2000; 2000US-0244677P.
PR
XX 01-NOV-2000; 2000US-0244984P.
PR
XX (FRAS/) FRASER M J.
PA (LIXX/) LI X.
PA (BEAM/) BEAM T.
PA (HUA/) HUA-VAN A.
XX
PI Fraser MJ, Li X, Beam T, Hua-Van A;
XX
XX MPI; 2003-352597/33.
DR
XX P-PSDB; ABU70357, ABU70358.
XX
XX New DNA molecule in the transposon piggyBac, useful for transferring
PT genes into host cells or embryos for transforming the cells of embryos
PT that can be used in making transgenic organisms.
PT
XX Disclosure; Fig 10(B); 151pp; English.
PS
XX The invention describes a DNA molecule comprising at least 163
CC consecutive nucleotide base pairs of the 3' terminal region beginning at
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of
CC the piggyBac molecule. The region extends from the restriction site SacI
CC to the end of the piggyBac molecule. The DNA molecule in the transposon
CC piggyBac is useful for transferring genes into host cells or embryos for
CC transforming the cells of embryos. The transposon contains genes for
CC useful for developing or making transgenic organisms. This sequence
CC represents plasmid used in the creation of minimal sequence eukaryotic
CC transformation vector piggyBac
XX
SQ Sequence 4613 BP; 1064 A; 1183 C; 1273 G; 1093 T; 0 U; 0 Other;

Query Match 100.0%; Score 707; DB 10; Length 4613;
Best Local Similarity 100.0%; Pred. No. 6.7e-213;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATCCCATCGCTCAATTTTACGACGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
DB 294 GGATCCCATCGCTCAATTTTACGACGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 353
QY 61 ATCATATCGTGGGTCTTTTTCGGGCTCAGTCAATCGCCCAAGCTGGCGTATCTGGGCA 120
DB 354 ATCATATCGTGGGTCTTTTTCGGGCTCAGTCAATCGCCCAAGCTGGCGTATCTGGGCA 413
QY 121 TCGGGGAGGAAGAGCCGCTCTTTTCCCGGAGTTCGAGCGGATGGAAGAGTTCG 180
DB 414 TCGGGGAGGAAGAGCCGCTCTTTTCCCGGAGTTCGAGCGGATGGAAGAGTTCG 473
QY 181 CCGAGGATGACTGCTGCTCATTGAGCGTTCGAGCGGATGGAAGAGTTCG 240
DB 474 CCGAGGATGACTGCTGCTCATTGAGCGTTCGAGCGGATGGAAGAGTTCG 533
QY 241 GGGAGGATGCTGGCATGCAATCGCTTTTAAACGGTGAATCTGTTTACGCGCACTGGGATA 300
DB 534 GGGAGGATGCTGGCATGCAATCGCTTTTAAACGGTGAATCTGTTTACGCGCACTGGGATA 593
QY 301 CCAGTTTCGTCGGGCTTTTTCGGGACACAGTTCGAGTTCGAGCGGATGGAAGAGTTCG 360
DB 594 CCAGTTTCGTCGGGCTTTTTCGGGACACAGTTCGAGTTCGAGCGGATGGAAGAGTTCG 653

Qy 361 ACCGGAACAATACCGCGCAGACGCGGAACTGCCGTCGCGTGCAGATTAAATGACAGCG 420
Db 654 ACCGGAACAATACCGCGCAGACGCGGAACTGCCGTCGCGTGCAGATTAAATGACAGCG 713
Qy 421 GTGCGGCGCTGGGATATTACGTTCAGCGAGCAGCGGTAATCCTGCTGATGCCGAGAAAT 480
Db 714 GTGCGGCGCTGGGATATTACGTTCAGCGAGCAGCGGTAATCCTGCTGATGCCGAGAAAT 773
Qy 481 GGACATGGATACCCCGTGAAGTTACCGCGGCGGCGCTCGTTCAATTCAGTTTGAAC 540
Db 774 GGACATGGATACCCCGTGAAGTTACCGCGGCGGCGCTCGTTCAATTCAGTTTGAAC 833
Qy 541 CCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTTTACAGGTGATGGAGCAGATGA 600
Db 834 CCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTTTACAGGTGATGGAGCAGATGA 893
Qy 601 AGATGCTCGACACGCTGCAGAACACGACGCTAGATTAAACCTAGAAAGATAATCATATTG 660
Db 894 AGATGCTCGACACGCTGCAGAACACGACGCTAGATTAAACCTAGAAAGATAATCATATTG 953
Qy 661 TGACGTACGTTAAAGATAATCATCGCTAAATTCAGCATGGGATCC 707
Db 954 TGACGTACGTTAAAGATAATCATCGCTAAATTCAGCATGGGATCC 1000

RESULT 3

ACA55361/c
ID ACA55361 standard; DNA; 8999 BP.

XX ACA55361;

DT 06-JUN-2003 (first entry)

XX Transformation vector piggyBAC related plasmid p(PZ)-Bac-EYFP.

XX PiggyBac; transposon; eukaryotic transformation vector; ds;

KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;

KW circular.

XX Synthetic.

XX US2002173634-A1.

XX 21-NOV-2002.

XX 30-OCT-2001; 2001US-00001189.

XX 31-OCT-2000; 2000US-0244677P.

PR 01-NOV-2000; 2000US-0244984P.

XX (FRAS/) FRASER M J.

PA (LIXX/) LI X.

PA (BEAM/) BEAM T.

PA (HUAU/) HUA-VAN A.

XX Fraser MJ, Li X, Beam T, Hua-Van A;

XX WPI; 2003-352597/33.

XX Example 11; Fig 12(B); 151pp; English.

CC The invention describes a DNA molecule comprising at least 163
CC consecutive nucleotide base pairs of the 3' terminal region beginning at
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of
CC the piggyBac molecule. The region extends from the restriction site SacI
CC to the end of the piggyBac molecule. The DNA molecule in the transposon
CC piggyBac is useful for transferring genes into host cells or embryos for
CC transforming the cells of embryos. The transformed cells or embryos are

CC useful for developing or making transgenic organisms. This sequence
CC represents plasmid used in the creation of minimal sequence eukaryotic
CC transformation vector piggyBac

SQ Sequence 8999 BP; 2180 A; 2416 C; 2342 G; 2061 T; 0 U; 0 Other;

Query Match 99.8%; Score 705.4; DB 10; Length 8999;

Best Local Similarity 99.9%; Pred. No. 3e-212;

Matches 706; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGATCCCATCGCTCAATTTTACGCACACTATCTTTCTAGGGTTAACTAGCTCATCAGG 60

Db 8950 GGATCCCATCGCTCAATTTTACGCACACTATCTTTCTAGGGTTAACTAGCTCATCAGG 8891

Qy 61 ATCATATCGTCGGGTCTTTTTCGGCTCAGTCATGCCCAAGCTGGCGTATCTGGGCA 120

Db 8890 ATCATATCGTCGGGTCTTTTTCGGGTCACTATGCCCAAGCTGGCGTATCTGGGCA 8831

Qy 121 TCGGGAGGAAGAAGCCCGTCTTTTCCGCGAGGTGGAAGCGGATGGAAGAGTTTG 180

Db 8830 TCGGGAGGAAGAAGCCCGTCTTTTCCGCGAGGTGGAAGCGGATGGAAGAGTTTG 8771

Qy 181 CCGAGGATGACTGCTGCTCATTTGAGCGGAAACGACGTTTACCATCATGATTC 240

Db 8770 CCGAGGATGACTGCTGCTCATTTGAGCGGAAACGACGTTTACCATCATGATTC 8711

Qy 241 GGGAGGTGTGGCCATGCACGCTTTTAAACGCTGAACTGTTTCAGGCCACCTGGGATA 300

Db 8710 GGGAGGTGTGGCCATGCACGCTTTTAAACGCTGAACTGTTTCAGGCCACCTGGGATA 8651

Qy 301 CCAGTTCTGCGGGCTTTTCCGGACACAGTTCGGATGTTCCGCCGAGCGCATCAGCA 360

Db 8650 CCAGTTCTGCGGGCTTTTCCGGACACAGTTCGGATGTTCCGCCGAGCGCATCAGCA 8591

Qy 361 ACCGGAACAATACCGCGCAGACGCGGAACTGCCGTCGCGTGCAGATTAAATGACAGCG 420

Db 8590 ACCGGAACAATACCGCGCAGACGCGGAACTGCCGTCGCGTGCAGATTAAATGACAGCG 8531

Qy 421 GTGCGGCGCTGGGATATTACGTTCAGCGAGGACGGGTATCCTGCTGGATGCCGAGAAAT 480

Db 8530 GTGCGGCGCTGGGATATTACGTTCAGCGAGGACGGGTATCCTGCTGGATGCCGAGAAAT 8471

Qy 481 GGACATGGATACCCCGTGAAGTTACCCGGCGGGCGCGCTCGTTCAATTCAGTTTGAAC 540

Db 8470 GGACATGGATACCCCGTGAAGTTACCCGGCGGGCGCGCTCGTTCAATTCAGTTTGAAC 8411

Qy 541 CCGTGGAGGACGGGACAGCTCCGCGTGCAAAATGTTTACAGCGGTGATGGAGCAGATGA 600

Db 8410 CCGTGGAGGACGGGACAGCTCCGCGTGCAAAATGTTTACAGCGGTGATGGAGCAGATGA 8351

Qy 601 AGATGCTCGACACGCTGCAGAACACGACGCTAGATTAAACCTAGAAAGATAATCATATTG 660

Db 8350 AGATGCTCGACACGCTGCAGAACACGACGCTAGATTAAACCTAGAAAGATAATCATATTG 8291

Qy 661 TGACGTACGTTAAAGATAATCATCGCTAAATTCAGCATGGGATCC 707

Db 8290 TGACGTACGTTAAAGATAATCATCGCTAAATTCAGCATGGGATCC 8244

RESULT 4

ACA55362/c

ID ACA55362 standard; DNA; 9009 BP.

XX ACA55362;

XX 06-JUN-2003 (first entry)

DT Transformation vector piggyBAC related plasmid p(PZ)-Bac-ECFP.

XX PiggyBac; transposon; eukaryotic transformation vector; ds;

KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;

XX circular.

OS Synthetic.
XX US2002173634-A1.
XX 21-NOV-2002.
XX 30-OCT-2001; 2001US-00001189.
XX 31-OCT-2000; 2000US-0244677P.
XX 01-NOV-2000; 2000US-0244984P.
XX (FRAS/) FRASER M J.
XX (LIXX/) LI X.
XX (BEAM/) BEAM T.
XX (HUA/) HUA-VAN A.
XX Fraser MJ, Li X, Beam T, Hua-Van A;
XX WPI; 2003-352597/33.
XX New DNA molecule in the transposon piggyBac, useful for transferring
XX genes into host cells or embryos for transforming the cells of embryos
XX PT that can be used in making transgenic organisms.
XX Example 11; Fig 13(B); 151pp; English.
XX The invention describes a DNA molecule comprising at least 163
XX consecutive nucleotide base pairs of the 3' terminal region beginning at
XX the 3' terminal base pair, and at least 125 consecutive nucleotide base
XX pairs of the 5' terminal region beginning at the 5' terminal base pair, of
XX the piggyBac molecule. The region extends from the restriction site SacI
XX to the end of the piggyBac molecule. The DNA molecule in the transposon
XX piggyBac is useful for transferring genes into host cells or embryos for
XX transforming the cells of embryos. The transformed cells or embryos are
XX useful for developing or making transgenic organisms. This sequence
XX represents plasmid used in the creation of minimal sequence eukaryotic
XX transformation vector piggyBac
XX Sequence 9009 BP; 2179 A; 2429 C; 2341 G; 2058 T; 0 U; 2 Other;
Query Match 99.8%; Score 705.4; DB 10; Length 9009;
Best Local Similarity 99.9%; Pred. No. 3e-212;
Matches 706; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGATCCCATCGCTCAATTTTACGACAGACTATCTTCTAGGGTTAACTAGCTGCATCAGG 60
DB 8960 GGATCCCATCGCTCAATTTTACGACAGACTATCTTCTAGGGTTAACTAGCTGCATCAGG 8901
QY 61 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCAATCGCCCAAGCTGGCGCTATCTGGCA 120
DB 8900 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCAATCGCCCAAGCTGGCGCTATCTGGCA 8941
QY 121 TCGGGAGGAAGAAGCCCGTCCTTTTCCCGCAGGTTGAAGCGGATGGAAGAGTTTG 180
DB 8840 TCGGGAGGAAGAAGCCCGTCCTTTTCCCGCAGGTTGAAGCGGATGGAAGAGTTTG 8781
QY 181 CCGAGATGACTGCTGCTGATTCAGCTTTCAGCGAAGAACCGACGTTTACCATCATCATTC 240
DB 8780 CCGAGATGACTGCTGCTGATTCAGCTTTCAGCGAAGAACCGACGTTTACCATCATCATTC 8721
QY 241 GGGAAAGTGTGGCCATCAGCCCTTTAAACGGTGAACCTGTTTCAGGCCACCTGGGATA 300
DB 8720 GGGAAAGTGTGGCCATCAGCCCTTTAAACGGTGAACCTGTTTCAGGCCACCTGGGATA 8661
QY 301 CCAAGTTGCTCGGGCTTTTCCGACACAGTTCCGGATGTTGACCCGAGCGCATCAGCA 360
DB 8660 CCAAGTTGCTCGGGCTTTTCCGACACAGTTCCGGATGTTGACCCGAGCGCATCAGCA 8601
QY 361 ACCCGAACATACCGGACAGCCGGAACCTGCGCGTGTGCGAGATTAATGACAGCG 420
DB 8600 ACCCGAACATACCGGACAGCCGGAACCTGCGCGTGTGCGAGATTAATGACAGCG 8541
QY 421 GTGCGGCGCTGGGATATTACGTACGCGAGGACGGGTATCTGCTGGATGCGCGAGAAAT 480

DB 8540 GTGCGGCGCTGGGATATTACGTACGCGAGGACGGGTATCTGCTGGATGCGCGAGAAAT 8481
QY 481 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGGCTCGTTTCATTTCACGTTTTTGAAC 540
DB 8480 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGGCTCGTTTCATTTCACGTTTTTGAAC 8421
QY 541 CCGTGAGGACCGGCGAGACTCGCGGTGCAAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
DB 8420 CCGTGAGGACCGGCGAGACTCGCGGTGCAAAATGTGTTTTACAGCGTGATGGAGCAGATGA 8361
QY 601 AGATGCTCGACACGCTGCGAGAACACGACGCTAGATTAAACCTAGAAAGATAATCATATTG 660
DB 8360 AGATGCTCGACACGCTGCGAGAACACGACGCTAGATTAAACCTAGAAAGATAATCATATTG 8301
QY 661 TGAAGTACGTTAAAGNTAATCATGCTGTAATAATGAGCATGGATCC 707
DB 8300 TGAAGTACGTTAAAGNTAATCATGCTGTAATAATGAGCATGGATCC 8254
RESULT 5
ACM55363/C
ID ACM55363 standard; DNA; 9012 BP.
XX AC
XX ACA55363;
XX AC
XX 06-JUN-2003 (first entry)
XX Transformation vector piggyBAC related plasmid p(PZ)-Bac-BGFP.
XX PiggyBac; transposon; eukaryotic transformation vector; ds;
XX transformed cell; transformed embryo; transgenic; plasmid; cyclic;
XX circular.
XX Synthetic.
XX OS
XX US2002173634-A1.
XX 21-NOV-2002.
XX 30-OCT-2001; 2001US-00001189.
XX 31-OCT-2000; 2000US-0244677P.
XX 01-NOV-2000; 2000US-0244984P.
XX (FRAS/) FRASER M J.
XX (LIXX/) LI X.
XX (BEAM/) BEAM T.
XX (HUA/) HUA-VAN A.
XX Fraser MJ, Li X, Beam T, Hua-Van A;
XX WPI; 2003-352597/33.
XX New DNA molecule in the transposon piggyBac, useful for transferring
XX genes into host cells or embryos for transforming the cells of embryos
XX PT that can be used in making transgenic organisms.
XX Example 11; Fig 14(B); 151pp; English.
XX The invention describes a DNA molecule comprising at least 163
XX consecutive nucleotide base pairs of the 3' terminal region beginning at
XX the 3' terminal base pair, and at least 125 consecutive nucleotide base
XX pairs of the 5' terminal region beginning at the 5' terminal base pair, of
XX the piggyBac molecule. The region extends from the restriction site SacI
XX to the end of the piggyBac molecule. The DNA molecule in the transposon
XX piggyBac is useful for transferring genes into host cells or embryos for
XX transforming the cells of embryos. The transformed cells or embryos are
XX useful for developing or making transgenic organisms. This sequence
XX represents plasmid used in the creation of minimal sequence eukaryotic
XX transformation vector piggyBac
XX Sequence 9012 BP; 2180 A; 2419 C; 2353 G; 2060 T; 0 U; 0 Other;

Db	1199	GGACATGGATACCCCGTAGTTACCCGGCGG-----CTGGTCAITTCAGCTTTTTCGAAC	1252
Qy	541	CCGTGGAGGACGGGCAGACTCGCGGTGCAAAATGTGTTTTACGGCTGATGGAGCAGATGA	600
Db	1253	CCGTGGAGGACGGGCAGACTCGCGGTGCAAAATGTGTTTTACGGCTGATGGAGCAGATGA	1312
Qy	601	AGATGCTTCGACACGCTCGAGAACACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTG	660
Db	1313	AGATGCTTCGACACGCTCGAGAACACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTG	1372
Qy	661	TGACGTACGTTAAAGATAAATCATCGGTAAAAATTGACGCATGGGATCC	707
Db	1373	TGACGTACGTTAAAGATAAATCATCGGTAAAAATTGACGCATGGGATCC	1419

RESIT.T 7

ACA55367
ID ACA55367 standard; DNA: 4943 BP.

ACA55367;

06-JUN-2003 (first entry)

XX Transformation vector pGαVBC related plasmid pBS-ITR-ECFP.

XX
XX
KW pienvBac: transposon: eukaryotic transformation vector: ds:

KW transformed cell: transformed embryo: transgenic: plasmid: cyclic:
KW transformed cell: transformed embryo: transgenic: plasmid: cyclic:
KW transformed cell: transformed embryo: transgenic: plasmid: cyclic:

KW circular.

XX OS Synthetic.

AA
PN
US2002173634-A1.XX
PD 21-NOV-2002XX
PF 30-OCT-2001: 2001US-000001189.31-OCT-2000 2000IIS-0244677P
XX
PB

PR 01-NOV-2000; 2000US-0244984P.
XX

PA (FRAS/) FRASER M J.
PA (ITVY/) ITVY

PA	{LIXX/}	LI X.
PA	(BEAM/)	BEAM

PA (BEAM/) BEAM I.
PA (HUAV/) HUA-VAN

XX
 (HOUA//) HOUA-VAN A:

PI FRASER MU, LI X, BEALL I, HUA-VAN A,
XX

DR 2003-352591/33.
XX

PT New DNA molecule in the transposon piggyBac, useful for transferring
PT genes into host cells or embryos for transforming the cells of embryos

PT genes into host cells of embryos for transforming the cells of embryos that can be used in making transgenic organisms.

FI
XX
that can be used in making transgenic organisms.

PS
XX
Example 6: fig 18(B); 151pp; English.

The invention describes a DNA molecule comprising at least 1b3 consecutive nucleotide base pairs of the 3' terminal region beginning at the 3' terminal base pair, and at least 125 consecutive nucleotide base pairs of the 5' terminal region beginning at the 5' terminal base pair, of the piggyBac molecule. The region extends from the restriction site SacI to the end of the piggyBac molecule. The DNA molecule in the transposon piggyBac is useful for transferring genes into host cells or embryos for transforming the cells of embryos. The transformed cells or embryos are useful for developing or making transgenic organisms. This sequence represents plasmid used in the creation of minimal sequence eukaryotic transformation vector piggyBac

Sequence 4943 BP: 1221 A: 1305 C: 1245 G: 1172 T: 0 U: 0 Other: 0

Query Match 96.9%; Score 685; DB 10; Length 4943;
Best Local Similarity 99.2%; Pred. No. 6.6e-206;
Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

Qy	1	GGATCCATCGCTCGGTCCTTTTACGAGCATCTCTTTCTAGGGTTAACTAGCTGCATCAGG	60
Db	719	GGATCCCATCGCTCAATTTTACGAGCATCTCTTTCTAGGGTTAACTAGCTGCATCAGG	778
Qy	61	ATCATATCGTCCGGTCTTTTTTCCGGCTCAGTCAATCGCCCAAGCTGCGCTATCTGGGCA	120
Db	779	ATCATATCGTCCGGTCTTTTTTCCGGCTCAGTCAATCGCCCAAGCTGCGCTATCTGGCA	838
Qy	121	TCGGGAGGAGGAGCGCCGTGCTTTTCCCGGAGAGTTGAAGCGGCATGGAAGAGTTTG	180
Db	839	TCGGGAGGAGGAGCGCCGTGCTTTTCCCGGAGAGTTGAAGCGGCATGGAAGAGTTTG	898
Qy	181	CCGAGGATGACTGCTGCTGATTTGAGCTTCAGCGGAAACGACGTTTACCATGATGATTC	240
Db	899	CCGAGGATGACTGCTGCTGATTTGAGCTTCAGCGGAAACGACGTTTACCATGATGATTC	958
Qy	241	GGGAAGTGTTGGCATGCACGCTTTTAAACGGTGAACTGTGTTCAAGGCCACCTTGGGATA	300
Db	959	GGGAAGTGTTGGCATGCACGCTTTTAAACGGTGAACTGTGTTCAAGGCCACCTTGGGATA	1018
Qy	301	CCAGTTCGTGCGCGCTTTTCCGAGACAGTTCCGGAATGTCAGCCGGAAGCGATCAGCA	360
Db	1019	CCAGTTCGTGCGCGCTTTTCCGAGACAGTTCCGGAATGTCAGCCGGAAGCGATCAGCA	1078
Qy	361	ACCCGAACAATACCGGCAGACGCGGAACTGCGGTGTCGGTGTGCAGATTAATGACAGCG	420
Db	1079	ACCCGAACAATACCGGCAGACGCGGAACTGCGGTGTCGGTGTGCAGATTAATGACAGCG	1138
Qy	421	GTGCGGGCTGGGATATTACGTGACGAGAGACGGTATCCTGGCTGATGCCCGCAGAAAT	480
Db	1139	GTGCGGGCTGGGATATTACGTGACGAGAGACGGTATCCTGGCTGATGCCCGCAGAAAT	1198
Qy	481	GGACATGGATACCCCGTAGCTTACCGCGCGGCGCGCTCGTTTCATTTCAGCTTTTTCGAAC	540
Db	1199	GGACATGGATACCCCGTAGCTTACCCGCGG-----CTCGTTTCATTTCAGCTTTTTCGAAC	1252
Qy	541	CCGTGGAGGACGGGCAGACTGCGGGTGCAAAATGTGTTTTACGGGTGATCGGACGATGA	600
Db	1253	CCGTGGAGGACGGGCAGACTGCGGGTGCAAAATGTGTTTTACGGGTGATCGGACGATGA	1312
Qy	601	AGATGCTCGACAGCTGTCAGAAACGACAGCTAGATTAAACCTTAGAAGATAATCATATTG	660
Db	1313	AGATGCTCGACAGCTGTCAGAAACGACAGCTAGATTAAACCTTAGAAGATAATCATATTG	1372
Qy	661	TGACGTACGTTAAAGATAATCATCGGTAAATTTGACGCATGGGATCC	707
Db	1373	TGACGTACGTTAAAGATAATCATCGGTAAATTTGACGCATGGGATCC	1419

RESULT 8

ACA55368
ID ACA55368 standard: DNA: 4944 BP.

ACA55368:

XX
DT 06-JUN-2003 (first entry)

XX Transformation vector pBS-ITR-EGFP.

XX PiggyBac; transposon; eukaryotic transformation vector; ds;
KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;
KW circular

0.1 + 0.4 + 0.3 + 0.2 = 1.0

OS
YY

PN US2002173634-A1.
v v

PD 21-NOV-2002.

PF 30-OCT-2001; 2001US-00001189.

PR 31-OCT-2000; 2000US-0244677P.

PR 01-NOV-2000; 2000US-0244984P.

XX (FRAS/) FRASER M J.
PA (LIXX/) LI X.
PA (BEAM/) BEAM T.
PA (HUAV/) HUA-VAN A.
XX Fraser MJ, Li X, Beam T, Hua-Van A;
XX WPI; 2003-352597/33.
XX
XX New DNA molecule in the transposon piggyBac, useful for transferring
PT genes into host cells or embryos for transforming the cells of embryos
PT that can be used in making transgenic organisms.
XX
XX Example 6; Fig 19(B); 151pp; English.
XX
XX The invention describes a DNA molecule comprising at least 163
CC consecutive nucleotide base pairs of the 3' terminal region beginning at
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of
CC the piggyBac molecule. The region extends from the restriction site SacI
CC to the end of the piggyBac molecule. The DNA molecule in the transposon
CC piggyBac is useful for transferring genes into host cells or embryos for
CC transforming the cells of embryos. The transformed cells or embryos are
CC useful for developing or making transgenic organisms. This sequence
CC represents plasmid used in the creation of minimal sequence eukaryotic
CC transformation vector piggyBac
XX
XX Sequence 4944 BP; 1221 A; 1302 C; 1247 G; 1174 T; 0 U; 0 Other;
SQ
Query Match 96.4%; Score 681.8; DB 10; Length 4944;
Best Local Similarity 98.9%; Pred. No. 6.8e-205;
Matches 699; Conservative 0; Mismatches 2; Indels 6; Gaps 1;
QY 1 GGATCCCATCGTCGCAATTTTACGCAGACTATCTTTCTAGGGTTAACTAGTCGCATCAGG 60
DB 719 GGATCCCATCGTCGCAATTTTACGCAGACTATCTTTCTAGGGTTAACTAGTCGCATCAGG 778
QY 61 ATCATATCGTCGGGTCTTTTTCGGGTCAGTCATCGCCCAAGCTGGCGTATCTGGGCA 120
DB 779 ATCATATCGTCGGGTCTTTTTCGGGTCAGTCATCGCCCAAGCTGGCGTATCTGGGCA 838
QY 121 TCGGGGAGGAAGCCCGTCTTTTCCGCGAGGTGAAGCGGCATGGAAGAGTTTG 180
DB 839 TCGGGGAGGAAGCCCGTCTTTTCCGCGAGGTGAAGCGGCATGGAAGAGTTTG 898
QY 181 CCAGAGGATGACTGCTGCTGATTGACGTTTCAGCGGAAACGACGTTTACCATCATGATTC 240
DB 899 CCAGAGGATGACTGCTGCTGATTGACGTTTCAGCGGAAACGACGTTTACCATCATGATTC 958
QY 241 GGGAAAGGTGTGGCCATGCACGCTTTTAAACGGTGAACCTGTTTCAGGCCACCTGGGATA 300
DB 959 GGGAAAGGTGTGGCCATGCACGCTTTTAAACGGTGAACCTGTTTCAGGCCACCTGGGATA 1018
QY 301 CCAGTTGCTCGCGGCTTTTCCGAGACACAGTTCCGATGFTACCCCGAAGCGCATCAGCA 360
DB 1019 CCAGTTGCTCGCGGCTTTTCCGAGACACAGTTCCGATGFTACCCCGAAGCGCATCAGCA 1078
QY 361 ACCCGACATACCGGCGACGCGGAACCTGCGTCCGCTGTCAGATTAATGACAGCG 420
DB 1079 ACCCGACATACCGGCGACGCGGAACCTGCGTCCGCTGTCAGATTAATGACAGCG 1138
QY 421 GTGCGGCGCTGGGATATTAGCTCAGCGAGGACGGGTATCTGCTGATGATCCCGCAGAAAT 480
DB 1139 GTGCGGCGCTGGGATATTAGCTCAGCGAGGACGGGTATCTGCTGATGATCCCGCAGAAAT 1198
QY 481 GGACATGGATACCCCGTGAGTTTACCCGCGGCGCGCTCGTTCAATTCAGCTTTTGAAC 540
DB 1199 GGACATGGATACCCCGTGAGTTTACCCGCGGCGG-----CTCGTTCAATTCAGCTTTTGAAC 1252
QY 541 CCGTGGAGGACGGGACAGCTCGGGTGCRAATGTGTTTACAGCGTGTGATGACAGATGA 600
DB 1253 CCGTGGAGGACGGGACAGCTCGGGTGCRAATGTGTTTACAGCGTGTGATGACAGATGA 1312

QY 601 AGATGCTCGACACGCTGCGAGAACACGCGCTAGATTAACTAGAGAAAGATAATCATATTG 660
DB 1313 AGATGCTCGACACGCTGCGAGAACACGCGCTAGATTAACTAGAGAAAGATAATCATATTG 1372
QY 661 TGACGCTACGTTAAAGATAATCATGCGTAAATTTGACGCGATGGATCC 707
DB 1373 TGACGCTACGTTAAAGATAATCATGCGTAAATTTGACGCGATGGATCC 1419
RESULT 9
ADP70043
ID ADP70043 standard; DNA; 48502 BP.
XX AC ADP70043;
XX DT 26-AUG-2004 (first entry)
XX DE Nucleic acid detection-related bacteriophage lambda DNA SeqID6.
XX KW nucleic acid detection; nucleic acid amplification; microchip;
XX KW Isothermal; chip substrate; ds.
XX OS Bacteriophage lambda.
XX PN JP2004154008-A.
XX PD 03-JUN-2004.
XX PF 01-NOV-2002; 2002JP-00320330.
XX PR 01-NOV-2002; 2002JP-00320330.
XX PA (EIKE) EIKEN KAGAKU KK.
XX WPI; 2004-445081/42.
XX PT Detecting nucleic acid such as DNA, cDNA and RNA, involves performing
XX PT nucleic acid amplification in microchip under isothermal conditions, and
XX PT analyzing obtained amplified product on microchip.
XX PS Example 2; SEQ ID NO 6; 49pp; Japanese.
XX CC This invention relates to a novel method of detecting nucleic acid which
XX CC comprises performing nucleic acid amplification in a microchip under
XX CC isothermal conditions, and analysing the obtained amplified product on
XX CC the microchip. The method is useful for detecting nucleic acid and
XX CC enables performing a nucleic acid amplification reaction and its analysis
XX CC continuously on the microchip without affecting the chip substrate. The
XX CC method enables simple, reliable and cost-effective amplification,
XX CC analysis and detection of nucleic acid in a microchip. The invention
XX CC enables detection of trace amounts of nucleic acid with amplification
XX CC efficiency. The present sequence is that of a bacteriophage lambda DNA
XX CC sequence which was used in the exemplification of the invention.
SQ Sequence 48502 BP; 12334 A; 11362 C; 12820 G; 11986 T; 0 U; 0 Other;
Query Match 82.8%; Score 585.4; DB 12; Length 48502;
Best Local Similarity 99.8%; Pred. No. 8e-174;
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
DB 3055 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 3114
QY 105 TGGCGCTATCTGGGCATCGGGAGGAGAGCCCGTCCCTTTTCCGCGAGTTGAGCG 164
DB 3115 TGGCGCTATCTGGGCATCGGGAGGAGAGCCCGTCCCTTTTCCGCGAGTTGAGCG 3174
QY 165 GCATGAAAGAGTTTTCGCCGAGGATGACTCTGCTGATTGACGTTGAGGCGAAACGCGACG 224
DB 3175 GCATGAAAGAGTTTTCGCCGAGGATGACTCTGCTGATTGACGTTGAGGCGAAACGCGACG 3234

QY 225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCAGCGCTTTTAAACGGTGAACTGTTTCGTT 284
DB |||||
DB 3235 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCAGCGCTTTTAAACGGTGAACTGTTTCGTT 3294
QY 285 CAGGCCACCTGGGATACCAAGTTTCGTGCGCGCTTTTCCGAGACACAGTTCGGATGGTCAGC 344
DB |||||
DB 3295 CAGGCCACCTGGGATACCAAGTTTCGTGCGCGCTTTTCCGAGACACAGTTCGGATGGTCAGC 3354
QY 345 CCGAAGCGCATACGACACCGAACAATACCGGAGACAGCCGGAACATGCGGTGCGGTGTG 404
DB |||||
DB 3355 CCGAAGCGCATACGACACCGAACAATACCGGAGACAGCCGGAACATGCGGTGCGGTGTG 3414
QY 405 CAGATTAAATGACAGCGGTGCGGCGCTGGGATATTACGTACGAGGACAGCGGTATCTCTGGC 464
DB |||||
DB 3415 CAGATTAAATGACAGCGGTGCGGCGCTGGGATATTACGTACGAGGACAGCGGTATCTCTGGC 3474
QY 465 TGGATGCGCGAGAAATGGACATCGATACCCCGTGTAGTTTACCGCGCGGCGCGCTCGTTC 524
DB |||||
DB 3475 TGGATGCGCGAGAAATGGACATCGATACCCCGTGTAGTTTACCGCGCGGCGCGCTCGTTC 3534
QY 525 ATTACGTTTTTGAACCCGTTGGAGGACGGGACACTCGCGGTGCAAAATGTGTTTACAGC 584
DB |||||
DB 3535 ATTACGTTTTTGAACCCGTTGGAGGACGGGACACTCGCGGTGCAAAATGTGTTTACAGC 3594
QY 585 GTGATGGAGCAGATGAAGATGCTCGACACCGCTGCAGAACACGCGAGCT 631
DB |||||
DB 3595 GTGATGGAGCAGATGAAGATGCTCGACACCGCTGCAGAACACGCGAGCT 3641

RESULT 10
AAS44733
ID AAS44733 standard; DNA; 4164 BP.
XX AC AAS44733;
XX DT 18-DEC-2001 (first entry)
XX DE Bacteriophage lambda full-length polynucleotide sequence #158.
XX KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
KW nervous system disorder; inflammatory disorder; cell differentiation; ds;
KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
KW cytoskeletal; anirrhematic; antiarthritic; tendon; ligament; tissue repair;
KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
KW immunostimulant; analgesic; gene therapy.
XX OS Bacteriophage lambda.
XX PN WO200164834-A2.
XX PD 07-SEP-2001.
XX PF 26-FEB-2001; 2001WO-US004926.
XX PR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
PR 17-JUN-2000; 2000US-00597707.
PR 14-JUL-2000; 2000US-00616807.
PR 19-SEP-2000; 2000US-00664641.
XX (HYSE-) HYSEQ INC.
XX PA Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao Q, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
PI Drmanac R;
XX WPI; 2001-589862/66.
DR P-PSDB; AAU27833.
XX

PT Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis, treatment of cancer,
PT neurological, inflammatory disorders and for use in arrays for detection.
XX
PS Claim 1; SEQ ID NO 158; 153pp; English.
XX
CC Sequences AAS44576-AAS44919 represent full-length polynucleotides and
CC contig polynucleotides encoding polypeptides of the invention. The DNA
CC and protein sequences are useful for the treatment, diagnosis and
CC prevention of various types of disorder in a mammalian subject such as a
CC human, dog, monkey, mouse, hamster or rat. The disorders include cancers
CC such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
CC as multiple sclerosis, connective tissue disease, rheumatoid arthritis,
CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
CC bowel disease. The sequences exhibit activity relating to angiogenesis,
CC cell proliferation, cell differentiation, stem cell growth factor,
CC activin or inhibin. Therefore, they can be used to manipulate stem cells
CC in culture to give rise to neuroepithelial cells that can be used to
CC augment or replace cells damaged by illness, accidental damage or genetic
CC disorders. The sequences may also be used for regeneration of bone,
CC cartilage, tendons and ligaments and in tissue repair and burn healing.
CC Note: Some sequences for this patent did not form part of the printed
CC specification, but were obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 4164 BP; 906 A; 1123 C; 1307 G; 828 T; 0 U; 0 Other;

Query Match 82.1%; Score 580.6; DB 4; Length 4164;
Best Local Similarity 99.3%; Pred No. 7.8e-173;
Matches 583; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 45 ATCTAGTCATCAGGATCATATCGTCGGGTCTTTTCCGGCTCAGTCATCGCCCAAGC 104
DB |||||
DB 178 ATCCAGTCGATCAGGATCATATCGTCGGGTCTTTTCCGGCTCAGTCATCGCCCAAGC 237
QY 105 TGGCGCTATCTGCGGATCGGGAGAGAGAGCCGCTGCTTTCCCGGAGTTCAAGG 164
DB |||||
DB 238 TGGCGCTATCTGCGGATCGGGAGAGAGAGCCGCTGCTTTTCCCGGAGTTCAAGG 297
QY 165 GCATGAAAGAGCTTTTCCGAGGATCACTGCTCTCATTTGAGCTTCAGCGAAACGACG 224
DB |||||
DB 298 GCATGAAAGAGCTTTTCCGAGGATCACTGCTCTCATTTGAGCTTCAGCGAAACGACG 357
QY 225 TTATCATGATGATTCGGGAAGGTGTGGCCATGACCGCTTTTAAACGGTGAACTGTTTCGTT 284
DB |||||
DB 358 TTATCATGATGATTCGGGAAGGTGTGGCCATGACCGCTTTTAAACGGTGAACTGTTTCGTT 417
QY 285 CAGGCCACCTGGGATACCAAGTTTCGTGCGGCTTTTCCGAGACACAGTTCGGATGGTCAGC 344
DB |||||
DB 418 CAGGCCACCTGGGATACCAAGTTTCGTGCGGCTTTTCCGAGACACAGTTCGGATGGTCAGC 477
QY 345 CCGAAGCGCATACAGCAACCCGAAACAATACCGGCGACAGCCGGAACATGCGGTGCGGTGTG 404
DB |||||
DB 478 CCGAAGCGCATACAGCAACCCGAAACAATACCGGCGACAGCCGGAACATGCGGTGCGGTGTG 537
QY 405 CAGATTAAATGACAGCGGTGCGGCGCTGGGATATTACGTACGAGGACAGCGGTATCTCTGGC 464
DB |||||
DB 538 CAGATTAAATGACAGCGGTGCGGCGCTGGGATATTACGTACGAGGACAGCGGTATCTCTGGC 597
QY 465 TGGATGCGCGAGAAATGGACATGGATACCCCGTGTAGTTTACCGCGGCGCGCTCGTTC 524
DB |||||
DB 598 TATATGCGCGAGAAATGGACATGGATACCCCGTGTAGTTTACCGCGGCGCGCTCGTTC 657
QY 525 ATTACGTTTTTGAACCCGTTGGAGGACGGGACACTCGCGGTGCAAAATGTGTTTACAGC 584
DB |||||
DB 658 ATTACGTTTTTGAACCCGTTGGAGGACGGGACACTCGCGGTGCAAAATGTGTTTACAGC 717
QY 585 GTGATGGAGCAGATGAAGATGCTCGACACCGCTGCAGAACACGCGAGCT 631
DB |||||


```

Db      718 GTGATGGAGCAGATAAAGATGCTCGACACGCTCGAGAAACGCGAGCT 764
RESULT 11
ABS78924/c
ID      ABS78924 standard; DNA; 26565 BP.
XX
XX      ABS78924;
AC
DT      17-DEC-2002 (first entry)
XX
XX      E. coli CFT073 genomic sequence #91.
DE
XX      Pathogenic; Escherichia coli CFT073 infection; livestock; pyelonephritis;
KW      urinary tract infection; open reading frame; ORF; uropathogenic;
KW      antibacterial; aropathic; nephrotropic; gene; ds.
XX
XX      Escherichia coli.
OS
XX
XX      WO200259320-A2.
PN
XX
XX      01-AUG-2002.
PD
XX
XX      19-OCT-2001; 2001WO-US046833.
PF
XX
XX      19-OCT-2000; 2000US-0242412P.
PR
XX
XX      (WISC ) WISCONSIN ALUMNI RES FOUND.
PA
XX
XX      Blattner FR, Welch RA, Burland VD;
PI
XX
XX      WPI; 2002-691532/74.
DR
XX
XX      New DNA sequences of the pathogenic Escherichia coli CFT073 strain,
PT      useful for preventing or treating E. coli CFT073 infection in humans or
PT      livestock.
XX
XX      Claim 1; Page 276-291; 765pp; English.
PS
XX
XX      The present invention relates to polynucleotide sequences from the genome
CC      of the pathogenic Escherichia coli strain CFT073. Almost all the
CC      sequences present in E. coli CFT073 are absent in the previously
CC      sequenced laboratory strain K-12. The polynucleotide sequences of the
CC      invention are useful for preventing, diagnosing or treating E. coli
CC      CFT073 infection in humans or livestock. The polynucleotide sequences are
CC      useful for preventing urinary tract infections and pyelonephritis.
CC      Likewise, the polypeptides encoded by the different open reading frames
CC      (ORF1-5) are useful for generating a vaccine against uropathogenic E.
CC      coli strains. ABS78834-ABS79085 represent genomic sequences from E. coli
CC      strain CFT073
XX
XX      Sequence 26565 BP; 7042 A; 6727 C; 5778 G; 7016 T; 0 U; 2 Other;
SQ
Query Match      80.1%; Score 566.2; DB 6; Length 26565;
Best Local Similarity 97.8%; Pred. No. 7.2e-168;
Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy      45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
Db      3221 ATCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 3162
Qy      105 TGGCGCTATCTGGGCGATCGGGGAGGAAGAGCCCGTGCCTTTTCCCGGAGGTGAAGCG 164
Db      3161 TGGCGCTATCTGGGCGATCGGGGAGGAAGAGCCCGTGCCTTTTCCCGGAGGTGAAGCG 3102
Qy      165 GCATGGAAGAGTTTCCCGAGGATGACTGCTGCTGCAATTGACGTTGAGCGAAACGCGACG 224
Db      3101 GCATGGAAGAGTTTCCCGAGGAGGACTGTTGCTGCAATTGACGTTGAGCGAAACGCGACG 3042
Qy      225 TTATCCATGATGATTCGGGAAGGTGGCCATCGACGCTTTTAACCGTGAACCTGTTTCGTT 284
Db      3041 TTACCATGATGATTCGGGAAGGTGGCCATCGATCGCTTTTAACCGTGAACCTGTTTCGTT 2982

```

```

Query Match      80.1%; Score 566.2; DB 10; Length 26565;
Best Local Similarity 97.8%; Pred. No. 7.2e-168;
Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```



```

Db      15275 ATTACGCTCTTTGAACCCGTTGGAGGACGGGACACCCGGGTGCAAAATGTTTACAGC 15334
Qy      585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCGAAGAACACGCAGCT 631
Db      15335 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCGAAGAACACGCAGCT 15381

RESULT 14
ADC00585
ID      ADC00585 standard; DNA; 46897 BP.
XX
AC      ADC00585;
XX
XX      04-DEC-2003 (first entry)
XX
XX      Enterohaemorrhagic E. coli O157:H7-specific nucleic acid SEQ ID NO: 630.
XX
XX      ds; gene; enterohaemorrhagic; anti-bacterial.
XX
XX      Escherichia coli; O157:H7.
XX
XX      JP2002355074-A.
XX
XX      10-DEC-2002.
XX
XX      24-JAN-2002; 2002JP-00015959.
XX
XX      24-JAN-2001; 2001JP-00112010.
XX
XX      (UYTS-) UNIV TSUKUBA.
XX
XX      WPI; 2003-451640/43.
XX
XX      Enterohaemorrhagic Escherichia coli O157:H7-specific nucleic acid molecule
XX      and a polypeptide and its use, a polypeptide, a vector and a host cell.
XX
XX      Claim 2; SEQ ID NO 630; 2067pp; Japanese.
XX
XX      The invention relates to a novel enterohaemorrhagic Escherichia coli
XX      O157:H7-specific nucleic acid molecule. A polynucleotide of the invention
XX      has anti-bacterial activity. The polypeptide can be used in detection
XX      and/or treatment of O157:H7 infection. The nucleotide sequence of the
XX      genome of Enterohaemorrhagic E. coli O157:H7 was determined. The present
XX      sequence represents an E. coli O157:H7-specific nucleic acid of the
XX      invention.
XX
XX      Sequence 46897 BP; 11872 A; 10948 C; 12731 G; 11346 T; 0 U; 0 Other;
XX
XX      Query Match 79.9%; Score 564.6; DB 10; Length 46897;
XX      Best Local Similarity 97.8%; Pred. No. 3.le-167;
XX      Matches 573; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
XX
Qy      45 ATCTAGCTGCATCAGGATCATATCGTGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 104
Db      14795 ATCCAGCTGCATCAGATCATATCGTGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 14854
Qy      105 TGGCGCTATCTGGGCATCTGGGGAGGAAGACCGGTGCTTTTTCGGCGAGTTGAAGCG 164
Db      14855 TGGCGCTATCTGGGCATCTGGGGAGGAAGACCGGTGCTTTTTCGGCGAGTTGAAGCG 14914
Qy      165 GCATGGAACAGTTTCGGCGAGTACTGCTGCTGATTCAGTTGAGCGGAACGCACG 224
Db      14915 GCATGGAACAGTTTCCGGAAGATGACTGTGTCTGCAITGAGCGTGAACGCACG 14974
Qy      225 TTATACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCCTTTTAAACGGTGAACTGTTCTG 284
Db      14975 TTATACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCCTTTAAACGGTGAACTGTTCTG 15034
Qy      285 CAGGCACCTGGGATACCAAGTTTCGTCGGGGCTTTTCCGGAACAAGTTCGGATGTCAGC 344
Db      15035 CAGGCACCTGGGATACCGGTCCCTCGGACTGTTCCGGAACAAGTTCGGATGTCAGC 15094
Qy      285 CAGGCACCTGGGATACCAAGTTTCGTCGGGGCTTTTCCGGAACAAGTTCGGATGTCAGC 344

```

Qy	345	CCGAAGCGGATCAGCAACCCGGAACAATATACCGCGGACAGCCGGAACCTGCGGTGCTGC	404
Db	15095	CCGAAGCGGATCAGCAACCCGGAACAATATACCGCGGACAGCCGGAACCTGCGGTGCTGC	15154
Qy	405	CAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTCAGCGAGACGGGTATCCTGCGC	464
Db	15155	CAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTCAGCGAGACGGGTATCCTGCGC	15214
Qy	465	TGGATCGCGCAGAAAATGAGACATGGATATACCCCGTGAGTTACCCGCGGCGCGCTCGTTTC	524
Db	15215	TGGATCGCGCAGAAAATGAGACATGGATATACCCCGTGAGTTACCCGCGGCGCGCTCGTTTC	15274
Qy	525	ATTACAGTTTTTTGAACCCCGTGAGAGACGGGACAGATCAGCGGTGCAAAATGTTGTTTACAGC	584
Db	15275	ATTACAGTTTTTTGAACCCCGTGAGAGACGGGACAGATCAGCGGTGCAAAATGTTGTTTACAGC	15334
Qy	585	GTGATGGACAGATGAAGATGCTCGACACGCTGCAGACACGACGACGACGCT 631	
Db	15335	GTGATGGACAGATGAAGATGCTCGACACGCTGCAGACACGACGACGACGCT 15381	
RESULT 15			
ACD19238/C			
ID	ACD19238 standard; DNA; 22306 BP.		
XX	ACD19238;		
XX			
XX			
DT	27-OCT-2003 (revised)		
DT	21-AUG-2003 (first entry)		
XX	E. coli 0157 unique DNA sequence OZID_251.		
DE	OZID; ds; acute haemorrhagic colitis; haemolytic uraemic syndrome;		
KW	food poisoning.		
KW			
KW			
XX	Escherichia coli; strain 0157:H7.		
OS			
XX	US2003023075-A1.		
PN			
XX	30-JAN-2003.		
PD			
XX			
PF	01-APR-2002; 2002US-00114170.		
XX			
PR	04-DEC-1998; 98US-0110955P.		
PR	03-DEC-1999; 99US-00453702.		
XX	(BLAT/) BLATTNER P R.		
PA	(BURL/) BURLAND V D.		
PA	(PERN/) PERNA N T.		
PA	(PLUN/) PLUNKETT G.		
PA	(WELC/) WELCH R.		
XX			
PI	Blattner FR, Burland VD, Perna NT, Plunkett G, Welch R;		
DR			
XX	WPI; 2003-479497/45.		
DR			
PT	New DNA sequences from Escherichia coli strain 0157:H7, useful for		
PT	detecting E. coli 0157:H7 in a sample, or in designing diagnostic probes		
PT	which can be used to distinguish strain 0157:H7 from strain K12 using		
PT	molecular techniques.		
XX			
PS	Claim 16; SEQ ID NO 251; 33pp; English.		
XX			
CC	The invention relates to an isolated DNA molecule comprising an E. coli		
CC	strain 0157:H7 sequence selected from a clostridial cytotoxin-like gene,		
CC	a urasee gene cluster, a RTX toxin-like gene cluster, a locus of		
CC	enterocyte effacement and 2 genes from its associated lymphocytic phage		
CC	933W (a putative serine/threonine kinase and a tail fibre gene). E. coli		
CC	0157:H7 can cause food poisoning, specifically acute haemorrhagic colitis		
CC	(which can develop into haemolytic uraemic syndrome). Also included are		
CC	an isolated DNA molecule comprising a nucleotide sequence identical to at		
CC	least 25 contiguous nucleotides contained in DNA sequences selected from		
CC	ACD18988-ACD19242 (being 255 E.coli 0157 DNA sequences which are not		

CC found in E.coli K12), a recombinant DNA construction comprising the DNA
CC above and a method for detecting E. coli 0157:H7 (ATCC 43895) in a sample
CC (or distinguishing between 0157 and K12) using a probe derived from one
CC of the 255 sequences. The DNA sequences are useful in detecting E. coli
CC 0157:H7 in a sample, for the early diagnosis of humans and livestock
CC infected with 0157:H7, and in designing diagnostic probes which can be
CC used to distinguish strain 0157:H7 from strain K12 using molecular
CC techniques. The present sequence is one of the 255 E. coli 0157:H7 DNA
CC sequence (termed O21D.1-O21D255). Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030203075 (Updated on 27-OCT-2003
CC to standardise OS field)
XX
SQ

Sequence 22306 BP; 4744 A; 6842 C; 5512 G; 5192 T; 0 U; 16 Other;

Query Match 79.4%; Score 561.4; DB 9; Length 22306;
Best Local Similarity 97.3%; Pred. No. 2.2e-166;
Matches 571; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY	45	ATCTAGTCATCAGGATCATATCGTCGGTCTTTTTCGGCTCAGTCATCGCCCAAGC	104
Db	17193	ATCAGTTGATCAGGATCATATCGTCGGTCTTTTTCGGCTCAGTCATCGCCCAAGC	17134
QY	105	TGCGCTATCTGGGTCATCGGGAGGAAGCCCGTGCCTTTTCCCGGAGGTTGAAGCG	164
Db	17133	TGCGCTATCTGGGTCATCGGGAGGAAGCCCGTGCCTTTTCCCGGAGGTTGAAGCG	17074
QY	165	GCATGGAAGAGTTTCCGAGGATGACTGCTGTCATTGACGTTGAGCGAAGACGACG	224
Db	17073	GCATGGAAGAGTTTCCGAGGATGACTGCTGTCATTGACGTTGAGCGAAGACGACG	17014
QY	225	TTTACCATGATGATTCGGGAAGGTGGCCATCGACGCTTTTAAACGGTGAACCTGTT	284
Db	17013	TTTACCATGATGATTCGGGAAGGTGGCCATCGACGCTTTTAAACGGTGAACCTGTT	16954
QY	285	CAGGCCACCTGGGATACCAAGTTTCGTCGGCTTTTCCGGACACAGTTCGGATGGT	344
Db	16953	CAGGCCACCTGGGATACCAAGTTTCGTCGGCTTTTCCGGACACAGTTCGGATGGT	16894
QY	345	CCGAAGCGCATCAGCAACCCGGAACAATACCGGCGACAGCCGGAACTGCGGTG	404
Db	16893	CCGAAGCGCATCAGCAACCCGGAACAATACCGGCGACAGCCGGAACTGCGGTG	16834
QY	405	CAGATTAAATGACAGCGTGGCGCTGGGATATTACGTAGCGAGGACGGGTATCCT	464
Db	16833	CAGATTAAATGACAGCGTGGCGCTGGGATATTACGTAGCGAGGACGGGTATCCT	16774
QY	465	TGGATCCGCGAGAAATGGACATGGATACCCCGTGGTATACCCGGCGGCGCGCT	524
Db	16773	TGGATCCGCGAGAAATGGACATGGATACCCCGTGGTATACCCGGCGGCGCGCT	16714
QY	525	ATTACAGTTTGAACCCGTGGAGGACGGGACACTCGCGGTGCAATGTGTTTACAG	584
Db	16713	ATTACAGTTTGAACCCGTGGAGGACGGGACACTCGCGGTGCAATGTGTTTACAG	16654
QY	585	GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACCGAGCT	631
Db	16653	GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACCGAGCT	16607

Search completed: June 5, 2005, 14:16:33
Job time : 345.253 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 13:37:08 ; Search time 2963.77 Seconds
(without alignments)

11558.877 Million cell updates/sec

Title: US-10-826-523-40

Perfect score: 707

Sequence: 1 ggaatccatgcgtcaatttt.....aaattgaogcatgggatcc 707

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hlg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	585.4	82.8	20067	12 XXU02425	U02425 Cloning vec
2	585.4	82.8	20125	12 XXU02426	U02426 Cloning vec
3	585.4	82.8	42529	12 CVU39284	U39284 Cloning vec
4	585.4	82.8	42530	12 CVU39285	U39285 Cloning vec
5	585.4	82.8	42531	12 CVU39286	U39286 Cloning vec
6	585.4	82.8	42704	12 CVU37692	U37692 Cloning vec
7	585.4	82.8	48502	7 LAMCG	J02459 Bacterioph
8	574.6	81.3	5969	12 AF434923	AF434923 Expressio
9	574.6	81.3	6182	12 AF434925	AF434925 Expressio
10	574.6	81.3	6182	12 AF434926	AF434926 Expressio
11	574.6	81.3	6182	12 AF434927	AF434927 Expressio
12	574.6	81.3	6182	12 AF434931	AF434931 Expressio
13	574.6	81.3	6182	12 AF434932	AF434932 Expressio
14	574.6	81.3	7421	12 AF434924	AF434924 Expressio
15	574.6	81.3	7634	12 AF434928	AF434928 Expressio
16	574.6	81.3	7634	12 AF434929	AF434929 Expressio
17	574.6	81.3	7634	12 AF434930	AF434930 Expressio
18	574.6	81.3	7634	12 AF434933	AF434933 Expressio
19	574.6	81.3	7634	12 AF434934	AF434934 Expressio

20	566.2	80.1	14164	1	AE005330	Escherich
c 21	566.2	80.1	26565	6	AX702505	Sequence
22	566.2	80.1	46819	6	AR204176	Sequence
c 23	566.2	80.1	305325	1	AE016765	Escherich
24	564.6	79.9	46897	6	BD184772	Nucleic a
c 25	564.6	79.9	222605	1	AP002555	Escherich
26	564.6	79.4	22306	6	AR204355	Sequence
c 27	561.4	79.4	91772	6	BD184765	Nucleic a
c 28	561.4	79.4	286485	1	AP002557	Escherich
c 29	407	57.6	46375	7	AF064539	Bacterioph
c 30	395.8	56.0	64137	2	AC101160	Mus muscu
c 31	395.8	56.0	179840	2	AC068458	Homo sapi
c 32	395.8	56.0	179871	2	AC024602	Homo sapi
33	395.8	56.0	194028	2	AC111077	Homo sapi
34	395.8	56.0	206486	2	AC174218	Mus muscu
c 35	362.2	51.2	228516	2	AC008594	Homo sapi
c 36	343.2	48.5	68028	2	AC101473	Mus muscu
c 37	329.8	46.6	66450	2	AC060821	Homo sapi
38	320.4	45.3	63579	2	AC100018	Mus muscu
39	320.2	45.3	63690	2	AC125419	Homo sapi
c 40	231.8	32.8	29466	1	AE008818	Salmonell
41	223	31.5	8134	7	PH2HEADTL	M81255 Bacterioph
c 42	223	31.5	13264	1	AE005368	Escherich
43	223	31.5	13386	1	AE015100	Shigella
c 44	223	31.5	34063	6	AR204200	Sequence
c 45	223	31.5	44029	6	BD184774	Nucleic a

ALIGNMENTS

RESULT 1
XXU02425
LOCUS
DEFINITION Cloning vector lambda EMBL3, left arm.
ACCESSION U02425
VERSION U02425.1 GI:413791
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 20067)
AUTHORS Frischauf,A.M., Lehrach,H., Pouetka,A. and Murray,N.
TITLE Lambda replacement vectors carrying polylinker sequences
JOURNAL J. Mol. Biol. 170 (4), 827-842 (1983)
MEDLINE 84064856
PUBMED 6315951
REFERENCE 2 (bases 1 to 20067)
AUTHORS Kitts,P.A.
TITLE CLONTECH Vectors On Disc version 1.3
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 20067)
AUTHORS Kitts,P.A.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303, USA
COMMENT This sequence has been compiled from information in the sequence databases, published literature and other sources. If you suspect there is an error in this sequence, please contact CLONTECH's Technical Service Department at (415) 424-8222 or (800) 662-2566, extension 3 or E-mail TECH@CLONTECH.COM.

FEATURES
source
1..20067
/organism="Cloning vector lambda EMBL3"
/mol_type="genomic DNA"
/db_xref="taxon:31783"

ORIGIN

Query Match 82.8%; Score 585.4; DB 12; Length 20067;
Best Local Similarity 99.8%; Pred. No. 2.3e-148;
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 45 ATTAGTCGATCAGGATCATATCGTGGGTCCTTTTTCGGGCTAGTCATCCGCCAAGC 104


```

gene
complement(19507. .20760)
/gene="lacy"
complement(19507. .20760)
/gene="lacy"
/codon_start=1
/codon_table=1
/transl_table=11
/product="lactose permease"
/protein_id="AAC53648.1"
/db_xref="GI:1066305"
/translators="MYIKNTNFMFGLFFFFYFFIMGAYFFPPFIWLHDINHISKSD
TGIIFAAISLSFLQFLGSLDKLRKYLWITGLMLVMEAPFFIFTFGLPLOWV
ILVSGIVGGIYLCFCFNAGAPAEATIEKVRBSNFEFGARMGCGWALCAI1SH
MFTINNOFVWLSGSCALIIAIVLLFFAKTDPSGATVANAVGANHSAPSLKLAIEFR
QPKLFIUSLVYSGNACTYDFDQOFANFFTSFGATGQGVTRVYVITMGBELLNASIM
FFAPLIINIRIGKNALLACTIMSIRIIGSFATSLAEVILKTHMFEVYPLLVGCGF
KTIYSQVEVFSFATYIVLCFCFFKFLKALAMFMSVLAGNMYESIGFQAGAYLVGLVALGCF
TLISVFTLSQGGPLSLLRQVNEA"
complement(20812. .23859)
/gene="lacz"
complement(20812. .23859)
/gene="lacz"
/notes="EcoRI site normally present at the 3' end of lacz
has been removed by an uncharacterized mutation"
/codon_start=1
/transl_table=11
/product="beta-galactosidase"
/protein_id="AAC53649.1"
/db_xref="GI:1066306"
/translators="VVIQRDWNPGVTQLNRLAAHPFASWRNSEARTRDRPSQQLR
SLNGEWRFPANFPAEAVPESWLECDLPEADTVVVPNSWQMHGYDAPITNTVYPIIVN
PPVFTNPTGCTSLTNVDESMLQEGOTRIIDFVNSAFHLCWNGKRWGVTGQSDRPS
SEDFLSAFLRAGENLAVMLVRSWDSYLEDQWRMSGIFRDVSLHKKFTTQISDFH
VATRFNDDFRAVLEAVQCMGELRDYRVTVLSMQGETQVAGTAPFGGEIIDERGG
YADRVTLRAVENPKLWSAIIENLYRAVVELHTADGTIIIEAACDVGFRVRLTENGLL
LANKKPLIGVNRHEHHPHGVMDSEQTVQDIIILMKONNFNARCSHYPNHPLAYT
LCBRYGLIYVDEANIEITHGVPMNRLTDDRMLPAMSERVTRVWRDRNRPSPVILWSL
GNSHGAGVADRYMTIKSVDPSPVQEGGADTTAIDTICMWIARVDESDQFPAVP
KWIISKWLSNPSTPILICEAHAMGNSLGGFAKYQAQFPRYLQGGFVMDVQPS
LIKYDENGPSWAGYGGFDGTPNDRQFCMNLGVFADRTPHALTAKHQOQFPQFLRL
GOTLEVTSEYLFHRSDNELLHMVADQKPLASGEVPLDVAPOGKOLIEULPELQPS
AGQLTLTVRVQPNATSEAGHLSAQWRLAENLSVTLPAASHAIPHILTTSEMDPC
TELGKRWQNRSGFUSQWIGDKQLLTPLRQDFTRALDNDIGVSEATRIDPNANW
VERWKAAGYOAEALLOCTADTLADAVLITTAHMQHQKTLFI SRKTYRIDSGOM
AITVDVEASDTPHAPRIGLNCOLAQVAERVNMGLGPOENYPRDLTAAACFRDMDLPL
SDMYTVPVFSENGRCCTRELNYGPHQWRGDFQFNISRYSQOQLMETSHRLLHAE
GTWLNIDGFMHIGDSDSPSVSABEQLSAGRYHYQLVWCQK"
complement(23874. .23957)
/notes="multiple cloning site"
complement(23973. .24141)
/notes="rrnB1 terminator"
24475. .25665
/gene="tetR"
24475. .25665
/gene="tetK"
/codon_start=1
/transl_table=11
/product="tetracycline resistance protein"
/protein_id="AAC53650.1"
/db_xref="GI:1066307"
/translators="MKSNMALIVILGTVTVDVAVIGLVMPVLPGLLRDIVHSDSIASH
YGVLLALYALMQFLCAKNGALSDRFRRLVLLASLLGATIDYIAWATPVLVILVAG
IRVAGITGATGAVAGAYIADITDGEDRARHFGMSACFGVGMVAGPVAGGLGAI1SH
APFLAAAVLGNGLMGLSCFKGSGRPMPLRPNFVSSFRWAGMTI1VAALVIT
FFTMQLGVQPAALWTVIGDRFRWSATMIGLSLVFGLHAAQAQPVTPGATKRGCE
KQAILAGMAADLGYVLLATPATRCWMAFPIIMILLASGICGMALQMLSRQVDDDDHQG
QLQSGAALATSLTISITGFLVITAIYAAASATWGLAWIVGAALYVCLPALRRGAWSR
ATST"
33602. .35450
/notes="imm21 substitution"

misc_feature
terminator
gene
CDS

Query Match 82.8%; Score 585.4; DB 12; Length 42529;
Best Local Similarity 99.8%; Pred. No. 2.5e-148;

```

gene	complement (19507.. .20760)	Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CDS	/gene="lacY" complement (19507.. .20760) /gene="lacI" /codon_start=1 /transl_table=11 /product="lactose permease" /protein_id="AAC53651.1" /db_xref="GI:1086309" /translation="MYLKNTNMFGLPFFFFVFIMGAYFPFFPFIWLHDIHISKSD TGIIFAAISLFLQPLGLGLSLDLKRLKYLWLIITGLMVPAPFFIIFGLLQYN ILVGSIVGGYVGLFCFNAGAPAEATEKVSRRSEFEFGRMFGVCGWALCASIVGI MFTINQFVWLGGCALILAVLLFAKTDPSSATVANAGAHSAFSLKALELFR QPKLPLSLVIVGSCVVDVQDFQANFETSFPATGEQTRVFGYVTVWGLLNASIM FPAPLINRIGGNALLAGTINSVRIIGSSPATSALEVVILKTLHMFVPEVPLLVGCF LYITQFEVAFSAIIVLCFCFQQLAMIFMSVLGNMYESIGFQAYILVGLVALGF TLISVFTLSPGFLSLLRQNEVA" complement (20812.. .23859) /gene="lacZ" complement (20812.. .23859) /gene="lacZ" /note="EcoRI site normally present at the 3' end of lacZ has been removed by an uncharacterized mutation" /codon_start=1 /transl_table=11 /product="beta-galactosidase" /protein_id="AAC53652.1" /db_xref="GI:1086310" /translation="VVLQRDWNPGVTQLNLAAHPPFASWRNSBEATDRPSQQLR SLNGEWFAPFAPEASLECDLPEADTVVPSNQMHGYDAPITYNTVYPIITVN PPFVPTNPGCGSLTFNVDESLOQGTIRIIFDVNSAFPHLMCNGRWVYGQDSRLP SEFDLAFATGAGNRLAVMLRWSGYLEDDQDMRMSGIFRDVSLHLKPTTQISDFH VATFNDSPRAVLEAVOMCGELRDYLRVTVSLMOGETQVASTAPFGGEIIDERGG YADRVLRLNVPKLSAEIIPNLRYAVVELHTADGTLIEAECADVGFREVIENGLL LLMGKPLLRGNVRHEHPLHVGQDEQTNVQDILLKQNNFNARCSHYPNHPLYIT LCDRYGLYVDEANIEYHGMVPMNRLTDDPRWLPAMSERVTMVQDRNPSVYIWSL GNESGHGANHDALYRIKGVDPSPVQYEGGGADTTATDIICPMYARVDEDDPPFAPV KWSIKWSLSPGTRPLILCEYAHAMGNSLGFAYKWOAFROYPRYLRQGVFWDPVDQS LIXYDGNPWSYAGDGDTPNDQFCMNGLVFADRTPHALTEAKHQOQFQFELS AGQWLTVRVVQPNATSEAGHSIAWQWRLAENLSVTLPAASHAIPHLTSEMDFC IELGNKRWQPSGSLSQMWIGDKQLLTPTRDQFTRAPLDNDIIVSEATRIDGNAM VETKAAALHQAEALLOCTADTLADAVLITTAHAWHQOKTLFI SRKTVRIDGSGQM AITVDVEASDTPHAPRIGLNCOLAQVAERVNWMLGLPOENYPRDLTAACFDRDLDPL SDMYTPVFPSENGLRCTFELNGYBQHWGDFQENISRYSOQLMETSHRLLHAE GTWLNDGFMHGIGDSDWSPSAEFQLSAGRYHQLVWCQK" complement (23875.. .23958) /note="multiple cloning site" complement (23974.. .24142) /note="rrnBT1 terminator" 24476.. .25666 /gene="tetR" 24476.. .25666 /gene="tetR" /codon_start=1 /transl_table=11 /product="tetracycline resistance protein" /protein_id="AAC53653.1" /db_xref="GI:1086311" /translation="MKSNNALIVLGTVTLDVAVGILVMPVLPGLLRDIVHSDSIASH YGVLLALYALMQFLCAPVLGALSDRFREPVLLASLLGATIDYAIMATTPVLMWLYAG RIVAGITGATGAVAGYIADIIDGEDRARHFGLMSACFGVGMVGFVAGGLIGALSILH APFLAAVLNGLNLGGLQMOESHKGRPMPRAFNPVSSPRWARGMTIIVAAIMTV PFTIMLVQVPAALWIFGEDRFWSATMIGLSLAVFGLLHALAQAPVVGPAERGE KQAILAGALDGLVLLAFATRGWAFPMIILGASGIGMPALQALMSRQVDDDDHQG QLQGSAAATLSLTSITGLPVTIATYAAASATWNLGLAWI VGAALYLVCLPLRRGAWSR ATST"	QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTAGTCATCGCCCAAGC 104 DB 3055 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTAGTCATCGCCCAAGC 3114 QY 105 TGGCGCTATCTGGGCATCGGGGAGGAAGACCGTCCTTTTTCGGCGGAGGTTGAAGCG 164 DB 3115 TGGCGCTATCTGGGCATCGGGGAGGAAGACCGTCCTTTTTCGGCGGAGGTTGAAGCG 3174 QY 165 GCATGAAAAGAGTTTCCGAGGATGACTCTGCTGCTGCTTACGCTTGAAGCGAAACGACG 284 DB 3175 GCATGAAAAGAGTTTCCGAGGATGACTCTGCTGCTGCTTACGCTTGAAGCGAAACGACG 3234 QY 225 TTTTACCATGATGATTCGGGAAGGTGGCCATGACGACGCTTTTAAACGGTGAACCTGTTTCGTT 284 DB 3235 TTTTACCATGATGATTCGGGAAGGTGGCCATGACGACGCTTTTAAACGGTGAACCTGTTTCGTT 3294 QY 285 CAGGCCACCTGGGATACCACTGTCGCGGCTTTTTCGGGACACAGTTCGGGATGTTGTCAGC 344 DB 3295 CAGGCCACCTGGGATACCACTGTCGCGGCTTTTTCGGGACACAGTTCGGGATGTTGTCAGC 3354 QY 345 CCGAAGCGCATGACGAACCCGGAACAATACCGGGACAGCGGGAACCTGCGGTGCGGTGTG 404 DB 3355 CCGAAGCGCATGACGAACCCGGAACAATACCGGGACAGCGGGAACCTGCGGTGCGGTGTG 3414 QY 405 CAGATTAAATGACAGCGGTGCGGCTGGGATATTACGTGACGAGGACGGGTATCTCTGGC 464 DB 3415 CAGATTAAATGACAGCGGTGCGGCTGGGATATTACGTGACGAGGACGGGTATCTCTGGC 3474 QY 465 TGGATGCCCGCAGAAATGAGCATGATACCCCGTGAAGTTACCGCGGGCGCGCTCGTTC 524 DB 3475 TGGATGCCCGCAGAAATGAGCATGATGATACCCCGTGAAGTTACCGCGGGCGCGCTCGTTC 3534 QY 525 ATTACAGTTTTTGAACCCCTGGAGGACGGGACAGCTCGCGGTGCAAAATGTGTTTTACAGC 584 DB 3535 ATTACAGTTTTTGAACCCCTGGAGGACGGGACAGCTCGCGGTGCAAAATGTGTTTTACAGC 3594 QY 585 GTGATCGACGACATGAAGATGCTCGACACGCTGACAGAACACGCGCT 631 DB 3595 GTGATCGACGACATGAAGATGCTCGACACGCTGACAGAACACGCGCT 3641
misc_feature	RESULT 5 LOCUS CUU39286 42531 bp DNA linear SYN 13-APR-1996 DEFINITION Cloning vector TLF97-3, phage lambda lacZ translational fusion vector, complete sequence. ACCESSION U39286 VERSION U39286.1 GI:1066312 KEYWORDS Cloning vector TLF97-3 SOURCE Cloning vector TLF97-3 ORGANISM Cloning vector TLF97-3 REFERENCE 1 (bases 1 to 42531) AUTHORS St Pierre,R. and Linn,T. TITLE A refined vector system for the in vitro construction of single-copy transcriptional or translational fusions to lacZ JOURNAL Gene 169 (1), 65-68 (1996) MEDLINE 96186904 PUBMED 8635751 REFERENCE 2 (bases 1 to 42531) AUTHORS St Pierre,R. TITLE Direct Submission JOURNAL Submitted (24-OCT-1995) Thomas Linn, Microbiology and Immunology, University of Western Ontario, London, Ontario N6A5C1, Canada FEATURES source 1. .42531 /organism="Cloning vector TLF97-3" /mol_type="genomic DNA" /db_xref="taxon:43842" terminator complement (19404.. .19489) /note="rrnBT2 terminator"	
gene	complement (19507.. .20760) /gene="lacY" complement (19507.. .20760) /gene="lacI" /codon_start=1 /transl_table=11 /product="lactose permease" /protein_id="AAC53651.1" /db_xref="GI:1086309" /translation="MYLKNTNMFGLPFFFFVFIMGAYFPFFPFIWLHDIHISKSD TGIIFAAISLFLQPLGLGLSLDLKRLKYLWLIITGLMVPAPFFIIFGLLQYN ILVGSIVGGYVGLFCFNAGAPAEATEKVSRRSEFEFGRMFGVCGWALCASIVGI MFTINQFVWLGGCALILAVLLFAKTDPSSATVANAGAHSAFSLKALELFR QPKLPLSLVIVGSCVVDVQDFQANFETSFPATGEQTRVFGYVTVWGLLNASIM FPAPLINRIGGNALLAGTINSVRIIGSSPATSALEVVILKTLHMFVPEVPLLVGCF LYITQFEVAFSAIIVLCFCFQQLAMIFMSVLGNMYESIGFQAYILVGLVALGF TLISVFTLSPGFLSLLRQNEVA" complement (20812.. .23859) /gene="lacZ" complement (20812.. .23859) /gene="lacZ" /note="EcoRI site normally present at the 3' end of lacZ has been removed by an uncharacterized mutation" /codon_start=1 /transl_table=11 /product="beta-galactosidase" /protein_id="AAC53652.1" /db_xref="GI:1086310" /translation="VVLQRDWNPGVTQLNLAAHPPFASWRNSBEATDRPSQQLR SLNGEWFAPFAPEASLECDLPEADTVVPSNQMHGYDAPITYNTVYPIITVN PPFVPTNPGCGSLTFNVDESLOQGTIRIIFDVNSAFPHLMCNGRWVYGQDSRLP SEFDLAFATGAGNRLAVMLRWSGYLEDDQDMRMSGIFRDVSLHLKPTTQISDFH VATFNDSPRAVLEAVOMCGELRDYLRVTVSLMOGETQVASTAPFGGEIIDERGG YADRVLRLNVPKLSAEIIPNLRYAVVELHTADGTLIEAECADVGFREVIENGLL LLMGKPLLRGNVRHEHPLHVGQDEQTNVQDILLKQNNFNARCSHYPNHPLYIT LCDRYGLYVDEANIEYHGMVPMNRLTDDPRWLPAMSERVTMVQDRNPSVYIWSL GNESGHGANHDALYRIKGVDPSPVQYEGGGADTTATDIICPMYARVDEDDPPFAPV KWSIKWSLSPGTRPLILCEYAHAMGNSLGFAYKWOAFROYPRYLRQGVFWDPVDQS LIXYDGNPWSYAGDGDTPNDQFCMNGLVFADRTPHALTEAKHQOQFQFELS AGQWLTVRVVQPNATSEAGHSIAWQWRLAENLSVTLPAASHAIPHLTSEMDFC IELGNKRWQPSGSLSQMWIGDKQLLTPTRDQFTRAPLDNDIIVSEATRIDGNAM VETKAAALHQAEALLOCTADTLADAVLITTAHAWHQOKTLFI SRKTVRIDGSGQM AITVDVEASDTPHAPRIGLNCOLAQVAERVNWMLGLPOENYPRDLTAACFDRDLDPL SDMYTPVFPSENGLRCTFELNGYBQHWGDFQENISRYSOQLMETSHRLLHAE GTWLNDGFMHGIGDSDWSPSAEFQLSAGRYHQLVWCQK" complement (23875.. .23958) /note="multiple cloning site" complement (23974.. .24142) /note="rrnBT1 terminator" 24476.. .25666 /gene="tetR" 24476.. .25666 /gene="tetR" /codon_start=1 /transl_table=11 /product="tetracycline resistance protein" /protein_id="AAC53653.1" /db_xref="GI:1086311" /translation="MKSNNALIVLGTVTLDVAVGILVMPVLPGLLRDIVHSDSIASH YGVLLALYALMQFLCAPVLGALSDRFREPVLLASLLGATIDYAIMATTPVLMWLYAG RIVAGITGATGAVAGYIADIIDGEDRARHFGLMSACFGVGMVGFVAGGLIGALSILH APFLAAVLNGLNLGGLQMOESHKGRPMPRAFNPVSSPRWARGMTIIVAAIMTV PFTIMLVQVPAALWIFGEDRFWSATMIGLSLAVFGLLHALAQAPVVGPAERGE KQAILAGALDGLVLLAFATRGWAFPMIILGASGIGMPALQALMSRQVDDDDHQG QLQGSAAATLSLTSITGLPVTIATYAAASATWNLGLAWI VGAALYLVCLPLRRGAWSR ATST"	
misc_feature	33603.. .35451 /note="imm21 substitution"	
ORIGIN	82.8%; Score 585.4; DB 12; Length 42530; Query Match 99.8%; Pred. No. 2.5e-148;	
Best Local Similarity		


```
gene      complement(19507..20760)
/ gene="lacI"
CDS       complement(19507..20760)
/ gene="lacY"
/ codon_start=1
/ transl_table=11
/ product="lactose permease"
/ protein_id="AAC53654.1"
/ db_xref="GI:1066313"
/ translation="MYLKNTNFMFGLFFFFFFIFMGAYFFPPFIWLHDIHISKSD
TGIIFAAISLIFQLFGLSKLGRYLKLLMIITGLMVFAPFPIFFGPLLQVN
ILVSIIVGGIVLGFECFNAGAPAEAFTEKYSRRSNEPFRGARMGCVGMALCAIIVI
MFINQOFWLGSCALILAVLLFPAKTDPSSATVANAVGNSAFSLKLELFR
QPKLFLSLVIGVSCYIDVDFQDFNFFTSFFATGQGRVFGVYITVMEGLLNASIM
FFAPLIINRIGGKNALLAGTMSVRIIGSSFATSALEVLKTLHMFVFLLVQCF
KYTSPFVRSFATYILVCFCKQLAMIFMSVLAGNMYESIGFQGAIVLVLGLVGF
TLISVFTLSPGFLSLLRQNEVA"
/ complement(20812..23859)
/ gene="lacZ"
/ complement(20812..23859)
/ gene="lacZ"
/ note="EcoRI site normally present at the 3' end of lacZ
has been removed by an uncharacterized mutation"
/ codon_start=1
/ transl_table=11
/ product="beta-galactosidase"
/ protein_id="AAC53655.1"
/ db_xref="GI:1066314"
/ translation="VVLQRDWNPGVTQLRLAAHPFASWRNSBEARTDRPSQQLR
SLNGEWRPAPFAPEAVPESWLECDLEADTVVPSNMQHGVDAPITNTVPIYVN
PPFVPTNPTGCSLTIVDDESWLQEGQRIIFDGVNSAFHLCNCRGWVGQDSRIP
SEFLSAFLGRNRLAVMVLRSQDSYLEDQWMSGIFRDVLSLLHKTPTQISDPH
VATRFNDPFLAVLEAEVQCELRDLRVTVLSMQGETQVASFAPFGGEIIDERGG
YADRTLRLNVEVPLKMSAEIPNLRYAVLVELTADGTLIEAEACDVGFREVRLENGLL
LNGKPLVIVGNRHEHPHVGQVMDQTVODILLKONNENAVRCSHPYHPHLAVT
LCRXYGLVYDEANLETHGQVPMNRLTDDPMLPAMSERVTRMVRQDRNHPSVLIWSL
GNSSGCHANDALYRWIKSDPSRPVQYEGGADTATDIIICPMYARVDEDDQPPAVP
KWSIKWLSLUPGETRPLICEYAHAMNSLGGFAKYQWAFPRYPLQGGVFWDMVDQS
LIKYDENGPWSAYGGDPTPNDRQFCMNLGVPADRTPALTEAKHQOFPQFLIS
GQTEIVTSEYLFPHSNEILLHVALIDGLASGEVPLDVAPQKOLIELPELPQES
AGOLWLTVRVQNPATASWAGHSIAWQWRLAENLSVTLPAASHAIPLHTTSEMDFC
IELGKWEQNRQSGLSOMWIGDKQLLPLRDQFTRAPLDNDIGVSEATRIDPNAM
VERKWAQGHQAEEALLOCTADTLADVLITTAHAWHQKQTLFISRKYTRIDGSGQM
AITVDVESADTTHPARGLNCLQAQVAERVNWMLGLGPQENYDRLTAAACFDRWDLPL
SDMYTPVFPFSGELGCRRELNYGHPQWRGDFQFNISRYSQQLMETSHRLLHABE
GTWNLIDGFMHIGGDSWSPSAEFLSAGRYHYQLVWCQR"
/ complement(23876..23959)
/ note="multiple cloning site"
/ notes="rrnBT1 terminator"
/ notes="rrnBT1 terminator"
/ gene="tetR"
/ db_xref="GI:1066315"
/ gene="tetR"
/ codon_start=1
/ transl_table=11
/ product="tetracycline resistance protein"
/ protein_id="AAC53656.1"
/ db_xref="GI:1066315"
/ translation="MKSNNALIVLGTVTLDAVGIGLVMPVLPGLLRDVIHSDSIASH
YGVLLALYALMQFLCAPVLGALSDFRGRPRVLLASLIGATIDYAIWATTVLWILYAG
RIVAGITGATGAVAGYIADITDGEDRARHFLMSACFCFVGVMVGPVAGGLGALSLH
APFLAALVNLGLNLLGCLFMQSHKERRPMPLRAFNPVSSFRWARGMTIVAAALTV
FFTMQLVGVQPAALWVTFGEDFRWSATMIGLSLAVFGILHALAQAPVTGPAFKRGE
KQAILAGMADALGYLVLAETRGWAFPMILLASGICMPALQMSLRQVDDDDHQG
QLQGSALALSLTISITGTLVITAIYASASTWNLGAWIVGAALYLVCLPALRRGANSR
ATST"
/ misc_feature 33604..33652
/ note="imm21 substitution"

ORIGIN
Query Match      82.8%; Score 585.4; DB 12; Length 42531;
Best Local Similarity 99.8%; Pred. No. 2.5e-148;
```

```
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 45 ATCTAGTCGATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCCGCCAAGC 104
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3055 ATCCAGTCGATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCCGCCAAGC 3114
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 105 TGGCGCTATCTGGGCATCGGGAGGAGAGCCCGTGCCTTTTCCCGCGAGGTTGAGCG 164
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3115 TGGCGCTATCTGGGCATCGGGAGGAGAGCCCGTGCCTTTTCCCGCGAGGTTGAGCG 3174
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 165 GCATGAAAGAGTTTCCCGAGGATGACTGCTGCTGCAATTCACCTTGAGCGAAGAACGACG 224
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3175 GCATGAAAGAGTTTCCCGAGGATGACTGCTGCTGCAATTCACCTTGAGCGAAGAACGACG 3234
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGACGACCTTTTAAACCGTGAACCTGTCGTT 284
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3235 TTTACCATGATGATTCGGGAAGGTGTGGCCATGACGACCTTTTAAACCGTGAACCTGTCGTT 3294
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 285 CAGGCCACCTGGGATACCACTTCGTCGGCTTTTCCGGACACAGTTCCCGATGGTCAGC 344
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3295 CAGGCCACCTGGGATACCACTTCGTCGGCTTTTCCGGACACAGTTCCCGATGGTCAGC 3354
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 345 CCGAAGCGCATCAGCAACCCGAAACATACCGGCGACAGCCGGAACCTGCCGTGCGGTGTG 404
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3355 CCGAAGCGCATCAGCAACCCGAAACATACCGGCGACAGCCGGAACCTGCCGTGCGGTGTG 3414
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 405 CAGATTAAATGACAGCGGTGCGGCGCTGGGATATTACGTACAGGAGGACGGGTATCCTGGC 464
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3415 CAGATTAAATGACAGCGGTGCGGCGCTGGGATATTACGTACAGGAGGACGGGTATCCTGGC 3474
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 465 TGATGCGCGAGAAATGGACATGATACCCGTGAGTTACCCGGCGGGCGCGCTCGTTC 524
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3475 TGGATGCGCGAGAAATGGACATGATACCCGTGAGTTACCCGGCGGGCGCGCTCGTTC 3534
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 525 ATTACAGTTTTTCAACCCGTGGAGGACGGCAGACTCGCGGTGCAAAATGTTTACAGC 584
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3535 ATTACAGTTTTTCAACCCGTGGAGGACGGCAGACTCGCGGTGCAAAATGTTTACAGC 3594
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 585 GTGATGAGCAGATGAAGATGCTCGACAGCTCGAGAACACGAGCT 631
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3595 GTGATGAGCAGATGAAGATGCTCGACAGCTCGAGAACACGAGCT 3641
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 6
LOCUS      CVU37692          42704 bp    DNA    linear    SYN 13-APR-1996
DEFINITION Cloning vector lambda TxF97, lacZ transcriptional fusion vector,
complete sequence.
ACCESSION  U37692
VERSION     U37692.1
KEYWORDS   GI:1051181
SOURCE     Cloning vector lambda TxF97
ORGANISM   Cloning vector lambda TxF97
REFERENCE  1 (bases 1 to 42704)
AUTHORS    St Pierre, R. and Linn, T.
TITLE      A refined vector system for the in vitro construction of
single-copy transcriptional or translational fusions to lacZ
JOURNAL    Gene 169 (1), 65-68 (1996)
MEDLINE    86186904
PUBMED     8635751
REFERENCE  2 (bases 1 to 42704)
AUTHORS    St Pierre, R. and Linn, T.
TITLE      Direct Submission
JOURNAL    Submitted (29-SEP-1995) Thomas Linn, Microbiology and Immunology,
University of Western Ontario, London, Ontario N6A5C1, Canada
FEATURES   Location/Qualifiers
            source          1..42704
                        /organism="Cloning vector lambda TxF97"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:43300"
            misc_feature    1..42704
                        /note="EcoRI site normally present at this site has been
```

```
removed by an uncharacterized mutation"
5505
/misc_feature
/notes="BamHI site normally present at this site has been
removed by an uncharacterized mutation"
complement(19404..19499)
/notes="rrnB2 terminator"
gene
complement(19507..20760)
/genes="lacY"
CDS
complement(19507..20760)
/genes="lacZ"
/codon_start=1
/transl_table=11
/product="lactose permease"
/protein_id="AAC53645.1"
/db_xref="GI:1051182"
/translation="MYLKNTEWMEGLPEFFYFFITMGAYVEFPFPLWLDINHISKSD
TGFIASISLFLPQLGSLDKLKYLLMIITGMLVMPAFPIIFGRLLOYN
ILVGSIVGGIYGFNCNAGAPAEIEKVSRRSNPEFGARPGFCWGALCAISVGI
QPLINNOFVWLGSCALILAVLLFAKTDAPSATVANAGHSAFSLKALELFR
MFKLWFLSYIVGVSCYVDVQDFANFTSPFATGEQGTGVFGYVTTMGELINASIM
FPAPLIIIRIGGNALLACTIMSRIIGSSPATSALEVVLKLTLMFVPLVGCFF
KYITSQEVREFSATILVCFCEFKQLAMIFMSVLGNMYESIGFQGYAILVLGLVGF
TLISVTLSGFGLSLRLRQVNEVA"
complement(20812..23886)
/genes="lacZ"
complement(20812..23886)
/genes="lacZ"
/notes="EcoRI site normally present at the end of lacZ has
been removed by an uncharacterized mutation"
/codon_start=1
/transl_table=11
/product="beta-galactosidase"
/protein_id="AAC53646.1"
/db_xref="GI:1051183"
/translation="MTMIDTSLAVLQRRDWNPGVTQLNLAHPFPASWRNSEAR
TDRPSQLSLNGEMFAPFPAEPFESWLECDLPEADTVVVPNSNMQHGVDAPLIY
NVTYPTVNPFTPTGTCYSLTFNVDSEWLQEGQTRIIIDGVNSAFHLCNGRWV
GYQDRLSEFDFLSAFRAGENLAVMLRWSDGSLYLEDQDMWRMSGI PRDVSLLHK
PTQISDFHVAFTNDFFSRVLEAEVQMGELRDYLRVTSLWQGETOVASGAPFG
GELIDRGGVADRVTLRVNPKLSAEI PNLYRAVVELHTADGTLIEAEACDVGFR
EVRLEGLLLKPLIRGVNRHEHPHGLHQVMDEQTMVQDILLMKQNFNRVCRSH
YPNHPWLUTCDRIGLYVVDENIETHGMVPMNRITDDPRWLPAMSERVTMVRQDRN
HPSVITWSLGNESGHGANHDALRWIKSVDPSPVQYEGGADTTADIIICPMYARVD
EDQFPAPKWSIKKWSLIPGEPTRPLILCEYAHAMGSLGSAKYQAFQYPRLOQG
FVWDVQDQSLIKYDENGPMWSAYGDDPTNDROFCMGNLVPADTPHPALTEAKHQ
QDFQLSQTLEVTSEYLFPSHDNELLHMVALDKPLAAGEVPLDVAPOGKOLIE
LPBQLPESAGQWLTVRVQPNATASEAGHISANQWRLEAENLSVTLPASHALIPH
LITSEMDFCIELGNKRQWFRQSLQOMWIGDKQLLTPLRDQTRAPLDNDIGVSE
ATRDPNMYERKAAHQYQAEALIQCTADTLADAVLITTAHAWQHOKTLFISRKT
YRIDGQMAITVDVEVADTPHPARIGLNCQLAQVAERVNMVLGLGPQNYPRDLTAA
CFDRWDLPLSDMTYVYVFPSENGLRCTRELNGPHQWRGDFQFNISRYSQOQLMETS
HRHLHABEGTWNIDGPHMGIGGDDSNPSVSAERQLSAGRHYHQLVWCQK"
complement(24049..24132)
/notes="multiple cloning site"
complement(24148..24316)
/notes="rrnB1 terminator"
gene
24650..25840
/genes="tetR"
CDS
24650..25840
/genes="tetR"
/notes="Description: tetracycline resistance gene;
tetracycline resistance protein"
/codon_start=1
/transl_table=11
/protein_id="AAC53647.1"
/db_xref="GI:1051184"
/translation="MKSNNALIVILGTVTLDVAVIGLVMPVLPLGLLRDIDVHSDSIASH
YGVLLALYALMQFLCAFRFRFVPLIASLIGATIDYAIMATTPVLWILYAG
RIVAGITGAVAGVADIITDGEDRARRHSLACFGVGVMPVAGMLLCAISLH
APFLAAVLGNLLGCTLMQESHKERRPMLRAFNPVSSFRWARGMTIVAAALMTV
FFIMQLVQVPALWTFGEDRFRWATMIGLSLAVFGLIHALAQAFVTPGPKRGE
QQAITAGAADALGVLLPATRWAFPIIMILLASGGIGMPALQVLPALRDRDDHQG
QLOGSLAALTSLSITGLPIVTAIYAASASTWNLAWIVGAALYLVCLFALRRGAWSR
```

```
ATST"
31747
/misc_feature
/notes="EcoRI site normally present at this site has been
removed by an uncharacterized mutation"
33777..35625
/notes="imm21 substitution"
39168
/misc_feature
/notes="EcoRI site normally present at this site has been
removed by an uncharacterized mutation"
ORIGIN
Query Match 82.8%; Score 585.4; DB 12; Length 42704;
Best Local Similarity 99.8%; Pred. No. 2.5e-148;
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
DB 3055 ATCCAGCTGCATCAGGATCATATCGTCGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3114
QY 105 TGGCGCTATCTGGGCATCGGGGAGGAAGACCCGTCCTTTTCCCGGAGGTTGAAGCG 164
DB 3115 TGGCGCTATCTGGGCATCGGGGAGGAAGACCCGTCCTTTTCCCGGAGGTTGAAGCG 3174
QY 165 GCATGGAAGAGTTTGGCCGAGGATGACTCTCTGCTGCAATTGACGTTGAGCGAAGAACCCACG 224
DB 3175 GCATGGAAGAGTTTGGCCGAGGATGACTCTCTGCTGCAATTGACGTTGAGCGAAGAACCCACG 3234
QY 225 TTATCCATCATGATTCGGGAGAGTGTCGATGACGCTTTTAAACGGTCAACTGTTTCGTT 284
DB 3235 TTATCCATCATGATTCGGGAGAGTGTCGATGACGCTTTTAAACGGTCAACTGTTTCGTT 3294
QY 285 CAGGCCACCTGGGATACCAAGTTTCGTCGGGCTTTTCCGGACACAGTTCGGGATGTTGTCAGC 344
DB 3295 CAGGCCACCTGGGATACCAAGTTTCGTCGGGCTTTTCCGGACACAGTTCGGGATGTTGTCAGC 3354
QY 345 CCGAAGCGCATCAGCAACCCGAAACAAATACCGGCGACAGCCGAACTGCCGTGCGGTGTG 404
DB 3355 CCGAAGCGCATCAGCAACCCGAAACAAATACCGGCGACAGCCGAACTGCCGTGCGGTGTG 3414
QY 405 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTAGTCAGCGAGGACGGGTATCTCTGCG 464
DB 3415 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTAGTCAGCGAGGACGGGTATCTCTGCG 3474
QY 465 TGGATGCCCGCAGAAATGGACATGATACCCCGTAGTTTACCCGGCGGGCGCGCTCTGTC 524
DB 3475 TGGATGCCCGCAGAAATGGACATGATACCCCGTAGTTTACCCGGCGGGCGCGCTCTGTC 3534
QY 525 ATTACCGTTTTTGAACCCGTGAGGACCGGCGAGACTCGCGGTGCAAAATGTTTACAGC 584
DB 3535 ATTACCGTTTTTGAACCCGTGAGGACCGGCGAGACTCGCGGTGCAAAATGTTTACAGC 3594
QY 585 GTGATGGACAGATGAAGATGCTCGACACGCTGCAGAACACCGAGCT 631
DB 3595 GTGATGGACAGATGAAGATGCTCGACACGCTGCAGAACACCGAGCT 3641
RESULT 7
LAMCG 48502 bp DNA circular PHG 17-APR-2002
LOCUS Bacteriophage lambda, complete genome.
DEFINITION J02459 M17233 M24325 V00636 X00906
ACCESSION J02459.1 GI:215104
VERSION DNA-binding protein; circular; coat protein; complete genome;
KEYWORDS origin of replication; repressor; unidentified reading frame.
SOURCE Bacteriophage lambda
ORGANISM Bacteriophage lambda
REFERENCE 1 (bases 1 to 12)
AUTHORS Wu, R. and Taylor, E.
TITLE Nucleotide sequence analysis of DNA. II. Complete nucleotide
sequence of the cohesive ends of bacteriophage lambda DNA
J. Mol. Biol. 57 (3), 491-511 (1971)
```

71209066
MEDLINE
PUBMED
REFERENCE
2 (bases 45493 to 45963)
AUTHORS
Imada,M. and Tsugita,A.
TITLE
Amino acid sequence of lambda phage endolysin
JOURNAL
Nature New Biol. 233, 230-231 (1971)
REFERENCE
3 (sites)
AUTHORS
Weigel,P.H., Englund,P.T., Murray,K. and Old,R.W.
TITLE
The 3'-terminal nucleotide sequences of bacteriophage lambda DNA
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 70 (4), 1151-1155 (1973)
MEDLINE
73215915
PUBMED
4515613
REFERENCE
4 (bases 38597 to 38672)
AUTHORS
Dahlberg,J.E. and Blattner,F.R.
TITLE
In vitro transcription products of lambda DNA: Nucleotide sequences
and regulatory sites
JOURNAL
(in) Fox,C.F. and Robinson,W.S. (Eds.),
VIRUS RESEARCH. PROCEEDINGS OF 1973 ICN-UCLA SYMPOSIUM: 533-544;
Academic Press, New York (1973)
REFERENCE
5 (bases 37945 to 38027)
AUTHORS
Maniatis,T., Ptashne,M., Backman,K., Kield,D., Flashman,S.,
Jeffrey,A. and Maurer,R.
TITLE
Recognition sequences of repressor and polymerase in the operators
of bacteriophage lambda
JOURNAL
Cell 5 (2), 109-113 (1975)
MEDLINE
75185528
PUBMED
1095210
REFERENCE
6 (bases 35583 to 35600)
AUTHORS
Kleid,D.G., Agarwal,K.L. and Khorana,H.G.
TITLE
The nucleotide sequence in the promoter region of the gene N in
bacteriophage lambda
JOURNAL
J. Biol. Chem. 250 (14), 5574-5582 (1975)
MEDLINE
75189495
PUBMED
167018
REFERENCE
7 (bases 35434 to 35618)
AUTHORS
Dahlberg,J.E. and Blattner,F.R.
TITLE
Sequence of the promoter-operator proximal region of the major
leftward RNA of bacteriophage lambda
JOURNAL
Nucleic Acids Res. 2 (9), 1441-1458 (1975)
MEDLINE
76031664
PUBMED
1178525
REFERENCE
8 (bases 37945 to 38018)
AUTHORS
Maniatis,T., Jeffrey,A. and Kleid,D.G.
TITLE
Nucleotide sequence of the rightward operator of phage lambda
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 72 (3), 1184-1188 (1975)
MEDLINE
75158212
PUBMED
1055375
REFERENCE
9 (bases 44588 to 44773)
AUTHORS
Sklar,J., Yot,P. and Weissman,S.M.
TITLE
Determination of genes, restriction sites, and DNA sequences
surrounding the 6S RNA template of bacteriophage lambda
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 72 (5), 1817-1821 (1975)
MEDLINE
75217847
PUBMED
1098044
REFERENCE
10 (bases 37905 to 37989)
AUTHORS
Walz,A., Pirrotta,V. and Ineichen,K.
TITLE
Lambda repressor regulates the switch between PR and P_{RM} promoters
JOURNAL
Nature 262 (5570), 665-669 (1976)
MEDLINE
76267718
PUBMED
958438
REFERENCE
11 (bases 37946 to 38039)
AUTHORS
Smith,G.R., Eisen,H., Reichardt,L. and Hedgepeth,J.
TITLE
Deletions of lambda phage locating a prm mutation within the
rightward operator
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 73 (3), 712-716 (1976)
MEDLINE
76152323
PUBMED
1062780
REFERENCE
12 (bases 35578 to 35667; 37903 to 38027)
AUTHORS
Ptashne,M., Backman,K., Humayun,M.Z., Jeffrey,A., Maurer,R.,
Meyer,B. and Sauer,R.T.
TITLE
Autoregulation and function of a repressor in bacteriophage lambda
JOURNAL
Science 194 (4261), 156-161 (1976)
MEDLINE
76271154

959843
REFERENCE
13 (bases 35578 to 35667)
AUTHORS
Humayun,Z., Jeffrey,A. and Ptashne,M.
TITLE
Completed DNA sequences and organization of repressor-binding sites
in the operators of phage lambda
JOURNAL
J. Mol. Biol. 112 (2), 265-277 (1977)
MEDLINE
77209970
PUBMED
875019
REFERENCE
14 (bases 38610 to 38732)
AUTHORS
Scherer,G., Hobom,G. and Kossel,H.
TITLE
DNA base sequence of the po promoter region of phage lambda
JOURNAL
Nature 265 (5590), 117-121 (1977)
MEDLINE
77100320
PUBMED
834253
REFERENCE
15 (bases 38041 to 38241)
AUTHORS
Roberts,T.M., Shimatake,H., Brady,C. and Rosenberg,M.
TITLE
Sequence of Cro gene of bacteriophage lambda
JOURNAL
Nature 270 (5634), 274-275 (1977)
MEDLINE
78071724
PUBMED
593347
REFERENCE
16 (bases 27616 to 28935)
AUTHORS
Davies,R.W., Schreier,P.H. and Buchel,D.E.
TITLE
Nucleotide sequence of the attachment site of coliphage lambda
JOURNAL
Nature 270 (5639), 757-760 (1977)
MEDLINE
78071823
PUBMED
593399
REFERENCE
17 (bases 37206 to 37263; 37914 to 37970)
AUTHORS
Humayun,Z.
TITLE
DNA sequence at the end of the CI gene in bacteriophage lambda
JOURNAL
Nucleic Acids Res. 4 (7), 2137-2143 (1977)
MEDLINE
78011659
PUBMED
909767
REFERENCE
18 (bases 27617 to 27934)
AUTHORS
Landy,A. and Ross,W.
TITLE
Viral integration and excision: structure of the lambda att sites
JOURNAL
Science 197 (4309), 1147-1160 (1977)
MEDLINE
77258934
PUBMED
331474
REFERENCE
19 (bases 39062 to 39170)
AUTHORS
Denniston-Thompson,K., Moore,D.D., Kruger,K.E., Furth,M.E. and
Blattner,F.R.
TITLE
Physical structure of the replication origin of bacteriophage
lambda
JOURNAL
Science 198 (4321), 1051-1056 (1977)
MEDLINE
78054731
PUBMED
929187
REFERENCE
20 (bases 44467 to 44807)
AUTHORS
Sklar,J.L.
TITLE
Structure and function of two regions of DNA controlling the
synthesis of prokaryotic RNAs
JOURNAL
Thesis (1977)
REFERENCE
21 (sites)
AUTHORS
Adhya,S. and Gottesman,M.
TITLE
Control of transcription termination
JOURNAL
Annu. Rev. Biochem. 47, 967-996 (1978)
MEDLINE
78234064
PUBMED
354508
REFERENCE
22 (bases 13 to 72; 48391 to 48502)
AUTHORS
Nichols,B.P. and Donelson,J.E.
TITLE
178-Nucleotide sequence surrounding the cos site of bacteriophage
lambda DNA
JOURNAL
J. Virol. 26 (2), 429-434 (1978)
MEDLINE
78197067
PUBMED
666898
REFERENCE
23 (bases 37938 to 38016; 35589 to 35666)
AUTHORS
Flashman,S.M.
TITLE
Mutational analysis of the operators of bacteriophage lambda
JOURNAL
Mol. Gen. Genet. 166 (1), 61-73 (1978)
MEDLINE
79114073
PUBMED
368570
REFERENCE
24 (bases 37990 to 38982)
AUTHORS
Schwarz,S., Scherer,G., Hobom,G. and Kossel,H.
TITLE
Nucleotide sequence of cro, cII and part of the O gene in phage

lambda DNA
JOURNAL Nature 272 (5652), 410-414 (1978)
MEDLINE 78135462
PUBMED 264238
REFERENCE 25 (bases 38212 to 38362)
AUTHORS Rosenberg, M., Court, D., Shimatake, H., Brady, C. and Wulff, D.L.
TITLE The relationship between function and DNA sequence in an interictronic regulatory region in phage lambda
JOURNAL Nature 272 (5652), 414-423 (1978)
MEDLINE 78135463
PUBMED 634366
REFERENCE 26 (bases 37224 to 37940)
AUTHORS Sauer, R.T.
TITLE DNA sequence of the bacteriophage gamma CI gene
JOURNAL Nature 276 (5685), 301-302 (1978)
MEDLINE 79053284
PUBMED 714163
REFERENCE 27 (bases 38597 to 39688)
AUTHORS Scherer, G.
TITLE Nucleotide sequence of the O gene and of the origin of replication in bacteriophage lambda DNA
JOURNAL Nucleic Acids Res. 5 (9), 3141-3156 (1978)
MEDLINE 79033241
PUBMED 704348
REFERENCE 28 (bases 29711 to 29811; 31043 to 31058)
AUTHORS Davies, R.W., Schreier, P.H. and Buchel, D.E.
TITLE Determination of the endpoints of partial deletion mutants of the attachment site of bacteriophage lambda by DNA sequencing
JOURNAL Nucleic Acids Res. 5 (9), 3209-3218 (1978)
MEDLINE 79033246
PUBMED 704352
REFERENCE 29 (bases 21661 to 31129)
AUTHORS Hoes, R.H. and Landy, A.
TITLE Structure of the lambda att sites generated by int-dependent deletions
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 75 (11), 5437-5441 (1978)
MEDLINE 79074833
PUBMED 364480
REFERENCE 30 (bases 39453 to 38500)

Query Match 82.8%; Score 585.4; DB 7; Length 48502;

Best Local Similarity 99.8%; Pred. No. 2.6e-148;

Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 45 ATCTAGTCATCAGGATCATATCGTCGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 104
DB |||||
DB 3055 ATCCAGTCGATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 3114
QY 105 TGGCGCTATCTGGGCATCGGGGAGGAAGACCGTCCTTTTCCCGCGAGGTTGAAGCG 164
DB |||||
DB 3115 TGGCGCTATCTGGGCATCGGGGAGGAAGACCGTCCTTTTCCCGCGAGGTTGAAGCG 3174
QY 165 GCATGAAAGAGTTTCCGAGGATGACTGCTGCTGATGACCTTGAAGCGAAACCGCAGC 224
DB |||||
DB 3175 GCATGAAAGAGTTTCCGAGGATGACTGCTGCTGATGACCTTGAAGCGAAACCGCAGC 3234
QY 225 TTTTACCATGATGATTCGGGAGGATGCGCATCGACCTTTTAAACGTTGAACGTTTCGTT 284
DB |||||
DB 3235 TTTTACCATGATGATTCGGGAGGATGCGCATCGACCTTTTAAACGTTGAACGTTTCGTT 3294
QY 285 CAGGCACCTCGGATACCACTTCGTTCGGCTTTTCCGACACAGTTCGGGATGTCAGC 344
DB |||||
DB 3295 CAGGCACCTCGGATACCACTTCGTTCGGCTTTTCCGACACAGTTCGGGATGTCAGC 3354
QY 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGCGCAGCGCGAACTGCGCGTGTG 404
DB |||||
DB 3355 CCGAAGCGCATCAGCAACCCGAAACAATACCGCGCAGCGCGAACTGCGCGTGTG 3414
QY 405 CAGATTAATGACAGCGGTTCGGCTCGGGATATTAGTCAGCGAGGACGGGTATCTCGC 464
DB |||||
DB 3415 CAGATTAATGACAGCGGTTCGGCTCGGGATATTAGTCAGCGAGGACGGGTATCTCGC 3474
QY 465 TGGATGCCGAGAAATGGACATGATGATACCCCGTGAGTTATCCCGCGGGCGCGCTCGTTC 524

Db 3475 TGGATGCCGAGAAATGGACATGATACCCCGTGAGTTATACCCGGCGCGCTCGTTC 3534
QY 525 ATTCAGCTTTTGAACCCGTGGAGACGGGAGACTCGCGGTGCAAAATGTTTACAGC 584
Db 3535 ATTCAGCTTTTGAACCCGTGGAGACGGGAGACTCGCGGTGCAAAATGTTTACAGC 3594
QY 585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCGAGAACACGACGCT 631
Db 3595 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCGAGAACACGACGCT 3641

RESULT 8

AF434923

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

enhancer

promoter

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS

misc_feature

CDS


```
QY 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGCGCACAGCCGGAATCTCCGTCGCGGTGTG 404
|
|
|
Db 4021 CCGAAGCGCATCAGCAACCCGAAACAATACCGCGCACAGCCGGAATCTCCGTCGCGGTGTG 4080
|
|
|
QY 405 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTAAGTCAGCGAGGACGGGTATCTCTGGC 464
|
|
|
Db 4081 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTAAGTCAGCGAGGACGGGTATCTCTGGC 4140
|
|
|
QY 465 TGGATGCGCGAGAAATGAGATGATACCCCGTGAGTTACCGCGCGCGCGCTCGTTTC 524
|
|
|
Db 4141 TGGATGCGCGAGAAATGAGATGATACCCCGTGAGTTACCGCGCGCGCGCTCGTTTC 4200
|
|
|
QY 525 ATTCACGTTTTTGAACCCGTTGGAGACCGGACAGCTCGCGGTGCAAAATGTTTTACAGC 584
|
|
|
Db 4201 ATTCACGTTTTTGAACCCGTTGGAGACCGGACAGCTCGCGGTGCAAAATGTTTTACAGC 4260
|
|
|
QY 585 GTGATGGAGCAGATGAAGATGCTCGACACCGCTCGAGAACAC 625
|
|
|
Db 4261 GTGATGGAGCAGATGAAGATGCTCGACACCGCTCGAGAACAC 4301
|
|
|
RESULT 10
AF434926
LOCUS
DEFINITION Expression vector 409-MUT, DNA circular SYN 04-APR-2002
ACCESSION AF434926
VERSION AF434926.1 GI:16904161
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE Expression vector 409-MUT
Expression vector 409-MUT
other sequences; artificial sequences; vectors.
Zieler,H. and Huynh,C.Q.
intron-dependent stimulation of marker gene expression in cultured
insect cells
Insect Mol. Biol. 11 (1), 87-95 (2002)
MEDLINE 21830665
PUBMED 11841506
REFERENCE
AUTHORS Zieler,H. and Huynh,C.Q.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of
Parasitic Diseases, National Institutes of Health, 4 Center Drive
MSC 0425, Bethesda, MD 20892-0425, USA
LOCATION/Qualifiers
FEATURES
source
1. .6182
/organism="Expression vector 409-MUT"
/mol_type="genomic DNA"
/db_xref="taxon:176968"
/notes="insect-specific"
97. .614
/notes="Autographa californica nuclear polyhedrosis virus
hr5 enhancer"
enhancer
615. .1201
/notes="Autographa californica nuclear polyhedrosis virus
immediate early promoter"
promoter
1266. 1398
/notes="artificial intron with mutated 3' splice site"
intron
1541. .3193
/codon_start=1
/product="firefly luciferase"
/protein_id="AAL30784.1"
/db_xref="GI:16904162"
CDS
/translation="MEDAKNIKGPAPFPYPLEDCTAGEQLHKMKRYALVPGTIATFD
AHLENIYAEYFEMSVRLAEMKRYGLNTHNRIVVCSNSLOQFMFVIGALFVAV
APANDIYNRELNSMNIQPTVVFVSKGLQKILNVOKKLPKIIIMDSKTDYQG
FSQMYTFVFNQIIPDAILSVVPHFGMFTTLGLVLCGRVLMYRFEELFVLR
LQYKQSLVLLPTLSPFAKSLIDKVDLSNHEIASGAPLSKEVGAIVAKRFLP
GIRQGLTETTSAILITEPDDKPGAVGVFFFAKVDLDTGKTLGVNQRELVCY
RGPWMSVYNNPEANALIDKWLHSDIAVDEHVVFFIVDRLSLILKYGYQVA
PAELESILQHNPIDAGVAGLPDAGELPAAVVVLHGHKNTKEIIVDYVASQVTT
AKKLRGVVFDVVPKGLTKLDARKIREILIKAKKGGSKL"
```

```
misc_feature 3359. .3464
/note="Heliothis virescens p63 chaperonin gene double
poly-adenylation sequence"
CDS
3381. .5241
/codon_start=1
/product="beta-lactamase"
/protein_id="AAL30785.1"
/db_xref="GI:16904163"
/translation="MSIQHPRVALIPFFPAFCLPVFAHPETLVKVKDAEDQLGARVGY
IELDLNSGKILSEFRPEERFPMWTFKVLICGAVLSRIDAGQEQLGRRIHGYQNDELVE
YSPVEKHLTDGMVRELCSAAITWSDNTANLLLTITGGPKELFTALFLHMGGDHVTRL
DWEPELNEAIPIRDERDITPVMAMATTLKLTITGELLTLASRQQLIDWMEADKAGPL
LRSLALPAGNFIAIDKSGAGRGSRGRIIAALGPDGKPSRIIVIVITGSGQATMDERNROIA
ETGASLIKHW"
5529. .5948
/note="CoIE1"
rep_origin
ORIGIN
Query Match 81.3%; Score 574.6; DB 12; Length 6182;
Best Local Similarity 99.3%; Pred. No. 1.8e-145;
Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 45 ATCTAGCTGCATCAGATCATATCGTCGGGTCTTTTTCCGGCTCAGTCATCGCCCAAGC 104
|
|
|
Db 3721 ATCCAGCTGCATCAGATCATATCGTCGGGTCTTTTTCCGGCTCAGTCATCGCCCAAGC 3780
|
|
|
QY 105 TGGCGCTATCTGGGCATCGGGAGGAGAGCCCGTGCCTTTTCCCGCAGGTTGAAGCG 164
|
|
|
Db 3781 TGGCGCTATCTGGGCATCGGGAGGAGAGCCCGTGCCTTTTCCCGCAGGTTGAAGCG 3840
|
|
|
QY 165 GCATGAAAAGAGTTTCCCGAGGATGACTCTCTGCTGATTCGATTGAGCTTGAGCGAAAACGACG 224
|
|
|
Db 3841 GCATGAAAAGAGTTTCCCGAGGATGACTCTCTGCTGATTCGATTGAGCGAAAACGACG 3900
|
|
|
QY 225 TTTACCATGATGATTCGGGAAGGTGTGCGCATGACAGCCCTTTTAAACGGTGAACCTGTTCTGTT 284
|
|
|
Db 3901 TTTACCATGATGATTCGGGAAGGTGTGCGCATGACAGCCCTTTTAAACGGTGAACCTGTTCTGTT 3960
|
|
|
QY 285 CAGGCCACCTGGGATACACAGTTCTGTCGGCGCTTTTCCGACACAGTTCCGGATGGTCAGC 344
|
|
|
Db 3961 CAGGCCACCTGGGATACACAGTTCTGTCGGCGCTTTTCCGACACAGTTCCGGATGGTCAGC 4020
|
|
|
QY 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGCGCACAGCCGGAATCTCCGTCGCGGTGTG 404
|
|
|
Db 4021 CCGAAGCGCATCAGCAACCCGAAACAATACCGCGCACAGCCGGAATCTCCGTCGCGGTGTG 4080
|
|
|
QY 405 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTAAGTCAGCGAGGACGGGTATCTCTGGC 464
|
|
|
Db 4081 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTAAGTCAGCGAGGACGGGTATCTCTGGC 4140
|
|
|
QY 465 TGGATGCGCGAGAAATGAGATGATACCCCGTGAGTTACCGCGCGCGCGCTCGTTTC 524
|
|
|
Db 4141 TGGATGCGCGAGAAATGAGATGATACCCCGTGAGTTACCGCGCGCGCGCTCGTTTC 4200
|
|
|
QY 525 ATTCAGTTTTTGAACCCCGTTGGAGACCGGACAGCTCGCGGTGCAAAATGTTTTACAGC 584
|
|
|
Db 4201 ATTCAGTTTTTGAACCCCGTTGGAGACCGGACAGCTCGCGGTGCAAAATGTTTTACAGC 4260
|
|
|
QY 585 GTGATGGAGCAGATGAAGATGCTCGACACCGCTCGAGAACAC 625
|
|
|
Db 4261 GTGATGGAGCAGATGAAGATGCTCGACACCGCTCGAGAACAC 4301
|
|
|
RESULT 11
AF434927
LOCUS
DEFINITION Expression vector 410-FOR, DNA circular SYN 04-APR-2002
ACCESSION AF434927
VERSION AF434927.1 GI:16904164
KEYWORDS
SOURCE Expression vector 410-FOR
ORGANISM Expression vector 410-FOR
other sequences; artificial sequences; vectors.
```

```
REFERENCE 1 (bases 1 to 6182)
AUTHORS Zieler,H. and Huynh,C.Q.
TITLE Intron-dependent stimulation of marker gene expression in cultured
insect cells
JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)
MEDLINE 21830665
PUBMED 11841506
REFERENCE 2 (bases 1 to 6182)
AUTHORS Zieler,H. and Huynh,C.Q.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of
Parasitic Diseases, National Institutes of Health, 4 Center Drive
MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES
source
    1..6182
        /organism="Expression vector 410-FOR"
        /mol_type="genomic DNA"
        /db_xref="taxon:176969"
        /note="insect-specific"
    97..614
        /note="Autographa californica nuclear polyhedrosis virus
        hr5 enhancer"
    615..1201
        /note="Autographa californica nuclear polyhedrosis virus
        immediate early promoter"
    1328..2980
        /codon_start=1
        /product="firefly luciferase"
        /protein_id="AAL30786.1"
        /db_xref="GI:16904165"
        /translation="MEDAKNIKGPAPPYPLEDGTAGEQLHKMKRYALVPGTIAFTD
        AHIEVNTIAYFEMSVRLAEAMKRYGLNTHRIIVCSNSLQPFVGLGALFIGVAV
        APANDIYNERELNSMNIQPTVVFVSKGLQKILNVOKLPIQKIIMDSKTDYQG
        FQSMYTFVSHLPFGNEYDFPESFDKDTKILIMNSGSGTGLPKGVALPHRTACVR
        FSHARDIFGNOIIPDTAILSWPFGHFGMTTGLVLCGFRVLMVREELFLRS
        LDQVKIOSALVPTLFSFEAKSTLIDKVDLSNHEIASGAPLSKEVGEAVAKRFLP
        GIRQYGLTETTSAILITPDGDKPGAVKVPFPEAKVVDLDTGKTLGVNQNGELCV
        RGPIMSGYVNNPENTNALIDKQWLSHSDIAYWDEDEHFFIYDRLKSLIKYKGYQVA
        PAELSLIILQHPNIFDAGVAGLPDDAGELPAAVYVLEHGKWTKEIEIVYVASQVTT
        AKLRGCVWFVDEVPKGLTKLDARKIREILIKAKGKSKL"
    3196..3328
        /note="artificial intron"
    3359..3464
        /note="Heliothis virescens p63 chaperonin gene double
        poly-adenylation sequence"
    4381..5241
        /codon_start=1
        /product="beta-lactamase"
        /protein_id="AAL30787.1"
        /db_xref="GI:16904166"
        /translation="MSIQHFRVALIPFFAFCFLVPFAHPETLVKVKDAEDQLGARVGY
        IELNLSGKILLESFRRPEFPMSTFKVLLCGAVLSRIDAGQQLGRRIRHYSQNDLVE
        YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTITIGPKELTAFLHNMGDHVTFL
        DRWEPELNEALPNDERTTWPMVAMATTLRKLTGELLTLASRQQLIDMWEADKVGCLP
        LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIIVYITGSGQTMDERRRQIA
        EIGASLIKHW"
    5529..5948
        /note="ColE1"
rep_origin
    81.3%; Score 574.6; DB 12; Length 6182;
    Best Local Similarity 99.3%; Pred. No. 1.8e-145;
    Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
ORIGIN
    45 ATCTAGCTGCATCAGGATCATATCGTCGGGCTTTTTTCGGGCTCAGTCATCGCCCAAGC 104
    |||
    3721 ATCCAGCTGCATCAGGATCATATCGTCGGGCTTTTTTCGGGCTCAGTCATCGCCCAAGC 3780
    |||
    105 TGGCGCTATCTGGGCATCGGGGAGGAAGACCCGCTGCTTTTCCCGCGAGGTTGAAGCG 164
    |||
    3781 TGGCGCTATCTGGGCATCGGGGAGGAAGACCCGCTGCTTTTCCCGCGAGGTTGAAGCG 3840
    |||
    165 GCATGGAAGAGTTTGGCGGAGGATGACTGTGCTGTCATTGACGTTGAGCGGAAACGCACG 224
```

```
Db 3841 GCATGGAAGAGTTTGGCGGAGGATGACTGTGCTGTCATTGACGTTGAGCGGAAACGCACG 3900
Qy 225 TTTACCATGATGATTCGGGGAAGGTGTGCCATGCACGCCCTTTTAAACCGTGAACTGTTGGTT 284
|||
Db 3901 TTTACCATGATGATTCGGGGAAGGTGTGCCATGCACGCCCTTTTAAACCGTGAACTGTTGGTT 3960
Qy 285 CAGGCCACCTCGGATACCACTGTCGCGGCTTTTTCGGGACACAGTTCGGATCGGTGAGC 344
|||
Db 3961 CAGGCCACCTCGGATACCACTGTCGCGGCTTTTTCGGGACACAGTTCGGATCGGTGAGC 4020
Qy 345 CCGAAGCGCATCAGCAACCCGGAACAATACCGGCGACAGCGGAACTCCCGTGCCTGGTGTG 404
|||
Db 4021 CCGAAGCGCATCAGCAACCCGGAACAATACCGGCGACAGCGGAACTCCCGTGCCTGGTGTG 4080
Qy 405 CAGATTAAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
|||
Db 4081 CAGATTAAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 4140
Qy 465 TGGATGCCGAGAAATGGACATGATACCCGCTGAGTATACCGCGGCGGCGGCTCGTTC 524
|||
Db 4141 TGGATGCCGAGAAATGGACATGATACCCGCTGAGTATACCGCGGCGGCGGCTCGTTC 4200
Qy 525 ATTACAGCTTTTGAACCCGCTGGAGGACGGGACAGACTCGCGTGCAAATGTGTTTACAGC 584
|||
Db 4201 ATTACAGCTTTTGAACCCGCTGGAGGACGGGACAGACTCGCGTGCAAATGTGTTTACAGC 4260
Qy 585 GTGATGAGCAGATGAAGATGCTCGACACGCTGCAGACCGCTGCAGAACAC 625
|||
Db 4261 GTGATGAGCAGATGAAGATGCTCGACACCGCTGCAGAACAC 4301
|||
RESULT 12
AF434931
LOCUS AF434931 6182 bp DNA circular SYN 04-APR-2002
DEFINITION Expression vector 409-REV, complete sequence.
ACCESSION AF434931
VERSION AF434931.1 GI:16904176
KEYWORDS Intron-dependent stimulation of marker gene expression in cultured
insect cells
SOURCE Expression vector 409-REV
ORGANISM Expression vector 409-REV
other sequences; artificial sequences; vectors.
REFERENCE 1 (bases 1 to 6182)
AUTHORS Zieler,H. and Huynh,C.Q.
TITLE Intron-dependent stimulation of marker gene expression in cultured
insect cells
JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)
MEDLINE 21830665
PUBMED 11841506
REFERENCE 2 (bases 1 to 6182)
AUTHORS Zieler,H. and Huynh,C.Q.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of
Parasitic Diseases, National Institutes of Health, 4 Center Drive
MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES
source
    1..6182
        /organism="Expression vector 409-REV"
        /mol_type="genomic DNA"
        /db_xref="taxon:176973"
        /note="insect-specific"
    97..614
        /note="Autographa californica nuclear polyhedrosis virus
        hr5 enhancer"
    615..1201
        /note="Autographa californica nuclear polyhedrosis virus
        immediate early promoter"
    complement (1236..1368)
        /note="artificial intron"
    1541..3193
        /codon_start=1
        /product="firefly luciferase"
        /protein_id="AAL30794.1"
```



```
/db_xref="GI:16904177"
/translation="MEDAKNIKKGPAPFPVPLEDGTAGEQLHKAMKRYALVPGTIAFTD
AHEVNITAEYFEMSVRLAEAMKRYGLNTHRIIVCSSENSLOFPMPVLGALFIGVAV
APANDIYNRELLSNMISQPTVVFVSKGLOKILNVQKKPLIIQKIIIMDSKTDYQG
FSQMYTFVTSHLPQPGNEVDVPESFDRDKTALIMNSGSGTGLPKGVALPHRTACVR
FSHARDPIFGNOIIPDTALISVVPFHGFMFTTGLYLCGFRVLMVYFEFELFLRS
LQDYKIQSALLVPTLFSFAKSLIDKYLNSLHEIASGAPLSKEVGAIVAKRHELCP
GIRQGYGLTETTSAILITPEGDDKPGAVGVVFFFEAKVVDLDGTGLVGNORGELCV
RGPWMSGVNPNPEATNALIDKDWLHSGDIAWDDEHFFIVDRKLSLIKYGQOVA
PAELSEILIQHPNIFDAGVAGLDDAGELPAAVVLEHGKMTTEKEIVDYVASQVTT
AKKRGVVVFVDEVPKGLTGKLDARKIREILIKAKKGGKSKL"
misc_feature
3359..3464
/notes="Heliothis virescens p63 chaperonin gene double
poly-adenylation sequence"
4381..5241
/codon_start=1
/product="beta-lactamase"
/protein_id="AAL30795.1"
/db_xref="GI:16904178"
/translation="MSIOHFRVALIPFEAFCLPVFAHPETLVKVKDAEDQLGARVGY
IELDLNSGKILSFRRPEFRPMWSTFKVLLCGAVLSRIDAGQELGRRHYSDNDLVE
YSPVTEKHLTDGWTRELCGAALTMSDNTAANLLLTIGPKELTAFLHMGDHTVRL
DRWEPELNAIPIFDERDTPMVAMATTLRLKLLTGLTLLASRQQLIDWMEADKVGPL
LRSLAPAGFIADKSGAGRGSRGIITALLGPDGPKFSRIIVYITVTSQATMDERNQIA
EIGASLIKHW"
rep_origin
5529..5948
/notes="Cole1"

ORIGIN
Query Match      81.3%; Score 574.6; DB 12; Length 6182;
Best Local Similarity 99.3%; Pred. No. 1.8e-145;
Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 45 ATCTAGCTGATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGC 104
|||
Db 3721 ATCCAGCTGATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGC 3780
|||

QY 105 TGGCGCTATCTGGGCATCGGGAGGAAGACCGCGCTTTTCCCGCAGGTTTGAAGCG 164
|||
Db 3781 TGGCGCTATCTGGGCATCGGGAGGAAGACCGCGCTTTTCCCGCAGGTTTGAAGCG 3840
|||

QY 165 GCATGAAAGAGTTTCCGAGGATGACTGCTGCTGATTGACGTTGAGCGAAACGACG 224
|||
Db 3841 GCATGAAAGAGTTTCCGAGGATGACTGCTGCTGATTGACGTTGAGCGAAACGACG 3900
|||

QY 225 TTACCATGATGATTGGGAAGGTGGCCATGCAAGCCTTTAACCGTGAACGTTCGTT 284
|||
Db 3901 TTACCATGATGATTGGGAAGGTGGCCATGCAAGCCTTTAACCGTGAACGTTCGTT 3960
|||

QY 285 CAGGCCACCTGGGATACGATCTGTCGGGCTTTTCCGGACACAGTTTCCGGATGGTCAGC 344
|||
Db 3961 CAGGCCACCTGGGATACGATCTGTCGGGCTTTTCCGGACACAGTTTCCGGATGGTCAGC 4020
|||

QY 345 CCGAAGCGCATCAGCAACCGAACAATACCGGCGACAGCGGAACCTCCGTGCCGGTGTG 404
|||
Db 4021 CCGAAGCGCATCAGCAACCGAACAATACCGGCGACAGCGGAACCTCCGTGCCGGTGTG 4080
|||

QY 405 CAGATTAATGACAGCGTGGCGGCTGGGATATTAAGTACGACAGACGGGTATCTTGGC 464
|||
Db 4081 CAGATTAATGACAGCGTGGCGGCTGGGATATTAAGTACGACAGACGGGTATCTTGGC 4140
|||

QY 465 TGGATGCCGAGAAATGGACATGGATACCCCGTGAGTTACCGCGGGCGCGCTCGTTC 524
|||
Db 4141 TGGATGCCGAGAAATGGACATGGATACCCCGTGAGTTACCGCGGGCGCGCTCGTTC 4200
|||

QY 525 ATTACGCTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAAATGTTTACAGC 584
|||
Db 4201 ATTACGCTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAAATGTTTACAGC 4260
|||

QY 585 GTGATGAGCAGATGAAGATGCTCGACACGCTCGAGACAC 625
|||
Db 4261 GTGATGAGCAGATGAAGATGCTCGACACGCTCGAGACAC 4301
|||
```

```
RESULT 13
AF434932
LOCUS AF434932 6182 bp DNA circular SYN 04-APR-2002
DEFINITION Expression vector 410-REV, complete sequence.
ACCESSION AF434932
VERSION AF434932.1 GI:16904179
KEYWORDS
SOURCE Expression vector 410-REV
ORGANISM Expression vector 410-REV
REFERENCE 1 (bases 1 to 6182)
AUTHORS Zieler H. and Huyh, C.Q.
TITLE Intron-dependent stimulation of marker gene expression in cultured
insect cells
JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)
MEDLINE 21830665
PUBMED 11841506
REFERENCE 2 (bases 1 to 6182)
AUTHORS Zieler H. and Huyh, C.Q.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of
Parasitic Diseases, National Institutes of Health, 4 Center Drive
MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES
source
1..6182
/organism="Expression vector 410-REV"
/mol_type="genomic DNA"
/db_xref="taxon:176974"
/notes="insect-specific"
97..614
/notes="Autographa californica nuclear polyhedrosis virus
hr5 enhancer"
615..1201
/notes="Autographa californica nuclear polyhedrosis virus
immediate early promoter"
1328..2980
/codon_start=1
/product="firefly luciferase"
/protein_id="AAL30796.1"
/db_xref="GI:16904180"
/translation="MEDAKNIKKGPAPFPVPLEDGTAGEQLHKAMKRYALVPGTIAFTD
AHEVNITAEYFEMSVRLAEAMKRYGLNTHRIIVCSSENSLOFPMPVLGALFIGVAV
APANDIYNRELLSNMISQPTVVFVSKGLOKILNVQKKPLIIQKIIIMDSKTDYQG
FSQMYTFVTSHLPQPGNEVDVPESFDRDKTALIMNSGSGTGLPKGVALPHRTACVR
FSHARDPIFGNOIIPDTALISVVPFHGFMFTTGLYLCGFRVLMVYFEFELFLRS
LQDYKIQSALLVPTLFSFAKSLIDKYLNSLHEIASGAPLSKEVGAIVAKRHELCP
GIRQGYGLTETTSAILITPEGDDKPGAVGVVFFFEAKVVDLDGTGLVGNORGELCV
RGPWMSGVNPNPEATNALIDKDWLHSGDIAWDDEHFFIVDRKLSLIKYGQOVA
PAELSEILIQHPNIFDAGVAGLDDAGELPAAVVLEHGKMTTEKEIVDYVASQVTT
AKKRGVVVFVDEVPKGLTGKLDARKIREILIKAKKGGKSKL"
complement (3166..3298)
/notes="artificial intron"
3359..3464
/notes="Heliothis virescens p63 chaperonin gene double
poly-adenylation sequence"
4381..5241
/codon_start=1
/product="beta-lactamase"
/protein_id="AAL30797.1"
/db_xref="GI:16904181"
/translation="MSIOHFRVALIPFEAFCLPVFAHPETLVKVKDAEDQLGARVGY
IELDLNSGKILSFRRPEFRPMWSTFKVLLCGAVLSRIDAGQELGRRHYSDNDLVE
YSPVTEKHLTDGWTRELCGAALTMSDNTAANLLLTIGPKELTAFLHMGDHTVRL
DRWEPELNAIPIFDERDTPMVAMATTLRLKLLTGLTLLASRQQLIDWMEADKVGPL
LRSLAPAGFIADKSGAGRGSRGIITALLGPDGPKFSRIIVYITVTSQATMDERNQIA
EIGASLIKHW"
5529..5948
/notes="Cole1"
rep_origin
```

ORIGIN
Query Match 81.3%; Score 574.6; DB 12; Length 6182;
Best Local Similarity 99.3%; Pred. No. 1.8e-145;

Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
Qy 45	ATCTAGCTGCATCAGGATCATATCGTCGGGCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
Db	
3721	ATCCAGCTGCATCAGGATCATATCGTCGGGCTTTTTCGGGCTCAGTCATCGCCCAAGC 3780
Qy 105	TGGCGCTATCTGGGCATCGGGGAGGAGAACCGCGTGCCTTTTCCCGGAGGTTGAAGCG 164
Db	
3781	TGGCGCTATCTGGGCATCGGGGAGGAGAACCGCGTGCCTTTTCCCGGAGGTTGAAGCG 3840
Qy 165	GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCTTACGTTGAGCGAGAAACGACG 224
Db	
3841	GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCTTACGTTGAGCGAGAAACGACG 3900
Qy 225	TTTACCATGATGATTCCGGAAGGTTGTCGCATGCACGCTTTTAAACGTTGAACTGTTT 284
Db	
3901	TTTACCATGATGATTCCGGAAGGTTGTCGCATGCACGCTTTTAAACGTTGAACTGTTT 3960
Qy 285	CAGGCCACCTGGGATACAGTTCTGCGGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
Db	
3961	CAGGCCACCTGGGATACAGTTCTGCGGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 4020
Qy 345	CCGAAGCGCATCAGCAACCCGAAACAATACCGCGACAGCGGAACTCGCGTCCGGGTG 404
Db	
4021	CCGAAGCGCATCAGCAACCCGAAACAATACCGCGACAGCGGAACTCGCGTCCGGGTG 4080
Qy 405	CAGATTAAATGACAGCGTGGCGCTGGGATATACGTCAGCGAGGACGGGTATCTCTGC 464
Db	
4081	CAGATTAAATGACAGCGTGGCGCTGGGATATACGTCAGCGAGGACGGGTATCTCTGC 4140
Qy 465	TGGATGCCGAGAAATGACATGATATCCCGTGAATTTACCGCGGGCGCGCTCGTTC 524
Db	
4141	TGGATGCCGAGAAATGACATGATATCCCGTGAATTTACCGCGGGCGCGCTCGTTC 4200
Qy 525	ATTACAGCTTTTGAACCGTGGAGCGGCGAGACTCGCGTGCAATGTGTTTACAGC 584
Db	
4201	ATTACAGCTTTTGAACCGTGGAGCGGCGAGACTCGCGTGCAATGTGTTTACAGC 4260
Qy 585	GTGATGAGCAGATGAAGATGCTGCACACGTCGAGAACAC 625
Db	
4261	GTGATGAGCAGATGAAGATGCTGCACACGTCGATGAAC 4301
RESULT 14	
AF434924	
LOCUS	
DEFINITION Expression vector pACTIN-LUC, complete sequence.	
ACCESSION AF434924	
VERSION AF434924.1 GI:16904155	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
FEATURES	
source	
promoter	

CDS	/note="Drosophila melanogaster actin5C promoter"	
	2780..4432	
	/codon_start=1	
	/product="firefly luciferase"	
	/protein_id="AAL30780.1"	
	/db_xref="GI:16904156"	
	/translation="MEDAKNIKGPAPFPLEDGTAGEQLHKAMKEYALVPGTIATPD	
	AHIEVNITAEYFEMSVRLAEAMRWKRYGLNTHRIIVVCSNSLQFFMVLGALFIGVAV	
	APANDYINERELLNSMNI SQPTVFSVSKGLKILNVQKLPITQIKIIMDSKTDYQG	
	FQMYTFVTSHLPFPGNEVDVPESDRDKTIALIMNSGSGTLPKGVLPHTACTVCR	
	FSHARDIPFGNQIIPDTAILSVVPHHFGMTTGLYLICGFRVLVMEFEELFRLR	
	LDYKIOSALLVPTLSPFAKSTLDKYDLNLHETASGAPLSKEVGEAAKREHLR	
	GIRQGGLETTSAILITPEGDDKPGNAVKVPFFPKAVYDLDTGKTLGVNQGELCV	
	RGPMISGVYNNPEATNALIDKDWLHSGDIAWDEDEHFFIVDRLSKLKIKYGIQVA	
	PALESILILQHPNIFDAGVAGLDDAGELPAAVVLEHGKTMTEKEIIVDYVASQVTT	
	AKKLRGVVFDVVPKGLTGKLDARKIREILIKAKKGKSKL"	
misc_feature	4598..4703	
	/note="Heliothis virescens p63 chaperonin gene double	
	poly-adenylation sequence"	
CDS	5620..6480	
	/codon_start=1	
	/product="beta-lactamase"	
	/protein_id="AAL30781.1"	
	/db_xref="GI:16904157"	
	/translation="MSIQHFRVALIPFFAFCPLVPFAHPETLVKVKDAEDQLGARVGY	
	IELDLSGKILLESFRPEERPMSTKSLVLRIDAGQLQGRIRRIHNSQNDLVE	
	YSPVTRKHTDGMTRELCSAAITMSDNTAANLLITIGPKELTAPLHNMGDHTLR	
	DRWEPELNEAI PNDRDITMPVAMATLRKLTGELLTLASRQQLIDMDEADKVAGPL	
	LRSAIPAGWFIADKSGAGERSGRIITAAALGPDGKPSRIVIYITGSGATWDERNRQIA	
	EIGASLIKHW"	
rep_origin	6768..7187	
	/note="ColE1"	
ORIGIN		
Query Match	81.3%;	Score 574.6; DB 12; Length 7421;
Best Local Similarity	99.3%;	Pred. No. 1.8e-145;
Matches 577;	Conservative 0;	Mismatches 4; Indels 0; Gaps 0;
QY	45	ATCTAGCTGCATCAGGATCATATCGTCGGGCTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
Db	4960	ATCCAGCTGCATCAGGATCATATCGTCGGGCTCTTTTTCGGGCTCAGTCATCGCCCAAGC 5019
QY	105	TGGCGCTATCTGGGCATTCGGGAGGAGAAAGCCCGTGCCTTTTCCCGGAGGTTGAAGCG 164
Db	5020	TGGCGCTATCTGGGCATTCGGGAGGAGAAAGCCCGTGCCTTTTCCCGGAGGTTGAAGCG 5079
QY	165	GCATGGAAGAGTTTCCCGAGGATGACTGCTGCTGCAATTCACGTTGAGCGGAAACCGACG 224
Db	5080	GCATGGAAGAGTTTCCCGAGGATGACTGCTGCTGCAATTCACGTTGAGCGGAAACCGACG 5139
QY	225	TTTACCATGATGATTCGGGAAGGTGGCCATCGACGCTTTTAAACGTTGAACCTGTTTCGTT 284
Db	5140	TTTACCATGATGATTCGGGAAGGTGGCCATCGACGCTTTTAAACGTTGAACCTGTTTCGTT 5199
QY	285	CAGGCCACCTGGGATACAGTTCTGTCGGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
Db	5200	CAGGCCACCTGGGATACAGTTCTGTCGGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 5259
QY	345	CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCGGAACTCGCGTCCGGGTGTG 404
Db	5260	CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCGGAACTCGCGTCCGGGTGTG 5319
QY	405	CAGATTAAATGACAGCGGTGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCTCTGGC 464
Db	5320	CAGATTAAATGACAGCGGTGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCTCTGGC 5379
QY	465	TGGATGCCGAGAAATGGAGATACCGCGTGAATTTACCGGCGGGCGCGCTCTCGTTTC 524
Db	5380	TGGATGCCGAGAAATGGAGATACCGCGTGAATTTACCGGCGGGCGCGCTCTCGTTTC 5439
QY	525	ATTACAGTCTTTTGAACCCGTGGAGGACGGGACAGCTCGCGGTGCAATGTGTTTACAGC 584
Db	5440	ATTACAGTCTTTTGAACCCGTGGAGGACGGGACAGCTCGCGGTGCAATGTGTTTACAGC 5499

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 13:39:19 ; Search time 117.806 Seconds
(without alignments)
9819.908 Million cell updates/sec

Title: US-10-826-523-40
Perfect score: 707
Sequence: 1 ggaatccatgcgtcaatttt.....aaaattgacgcatggatcc 707

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	566.2	80.1	46819	3	US-09-453-702B-72
C 2	561.4	79.4	22306	3	US-09-453-702B-251
C 3	223	31.5	34083	3	US-09-453-702B-96
C 4	73.8	10.4	5194	3	US-08-844-274-16
C 5	73.8	10.4	5194	4	US-09-598-421-16
C 6	73.4	10.4	5194	3	US-08-844-274-17
C 7	73.4	10.4	5194	4	US-09-598-421-17
8	67.8	9.6	5679	3	US-08-844-274-10
9	67.8	9.6	5679	4	US-09-598-421-10
10	67.8	9.6	6723	3	US-08-844-274-13
C 11	67.8	9.6	6723	3	US-08-844-274-14
C 12	67.8	9.6	6723	4	US-09-598-421-13
C 13	67.8	9.6	6723	4	US-09-598-421-14
C 14	67.8	9.6	7560	3	US-08-844-274-20
C 15	67.8	9.6	7560	4	US-09-598-421-20
C 16	67.8	9.6	9423	3	US-09-377-066-6
C 17	64.4	9.1	6448	3	US-08-844-274-15
C 18	64.4	9.1	6448	4	US-09-598-421-15
C 19	63.8	9.0	2476	3	US-08-844-274-11
C 20	63.8	9.0	2476	4	US-09-598-421-11
21	39.8	5.6	489	4	US-09-452-991A-15689
22	39.8	5.6	888	4	US-09-252-991A-15722
C 23	39.8	5.6	2001	4	US-09-252-991A-15722
C 24	39.8	5.6	2475	4	US-09-252-991A-15589
C 25	39.2	5.5	1734	6	US-09-252-991A-15758
C 26	39.2	5.5	1734	6	5352575-8
C 27	35.4	5.0	209210	4	US-09-949-016-15094

Patent No. 5352575
Patent No. 5352575
Sequence 15094, A

Sequence 3154, Ap
Sequence 2816, Ap
Sequence 2, Appli
Sequence 1, Appli
Sequence 7848, Ap
Sequence 1011, Ap
Sequence 1011, Ap
Sequence 1011, Ap
Sequence 1011, Ap
Sequence 1011, Ap
Sequence 785, App
Sequence 1230, Ap
Sequence 573, App
Sequence 1868, Ap
Sequence 17150, A
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-453-702B-72
; Sequence 72, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Nicole T.
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod

TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0

CURRENT APPLICATION DATA: US/09/453,702B
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998

ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017

TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166

INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 46819
TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 72:

US-09-453-702B-72
Query Match 80.1%; Score 566.2; DB 3; Length 46819;
Best Local Similarity 97.8%; Pred. No. 6.6e-176;

Matches	574;	Conservative	0;	Mismatches	13;	Indels	0;	Gaps	0;
Qy	45	ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCCGGCTCAGTCATCGCCCAAGC	104						
Db	14795	ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCCGGCTCAGTCATCGCCCAAGC	14854						
Qy	105	TGGCGCTATCTGGGCATCGGGAGGAAGAAGCCGTGCCTTTTCCGGAGAGTTGAACGG	164						
Db	14855	TGSCGCTATCTGGGCATCGGGAGGAAGAAGCCGTGCCTTTTCCGGAGAGTTGAACGG	14914						
Qy	165	GCATGAAAGAGTTTCCCGAGAGATGACTCTGCTGCTATTCAGCTTTGAGCGAAAACGCACG	224						
Db	14915	GCATGAAAGAGTTTCCCGAAGATGACTGCTGCTATTCAGCTTTGAGCGAAAACGCACG	14974						
Qy	225	TTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTTAAACGGTGAACGTTCGTT	284						
Db	14975	TTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTTAAACGGTGAACGTTCGTT	15034						
Qy	285	CAGGCCACCTGGGATACAGTTCTGTCGGGCTTTTTCGGACACAGTTTCGGATGGTCAAGC	344						
Db	15035	CAGGCCACCTGGGATACCCGCTCCCTCGCACCTGTTCCGGACACAGTTTCGGATGGTCAAGC	15094						
Qy	345	CCGAAGCGCATCAGCAACCCGAAACAAATACCGCGACAGCCGGAACCTGCCCTGCCGCTGTG	404						
Db	15095	CCGAAGCGCATCAGCAACCCGAAACAAATACCGCGACAGCCGGAACCTGCCCTGCCGCTGTG	15154						
Qy	405	CAGATTAAATACACAGCGGTTCGGCGCTGGGATATTACGTACAGCGAGACGGGTATCTCGGC	464						
Db	15155	CAGATTAAATACACAGCGGTTCGGCGCTGGGATATTACGTACAGCGAGACGGGTATCTCGGC	15214						
Qy	465	TGATCCGCGAGAAATGGACATGGATACCCGTGAGTTTACCCGGCGGGCGCCTCGTTC	524						
Db	15215	TGATCCGCGAGAAATGGACATGGATACCCGTGAGTTTACCCGGCGGGCGCCTCGTTC	15274						
Qy	525	ATTTCAGTTTTTTGAAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAAATGTTTTTACAGC	584						
Db	15275	ATTTCAGTCTTTTGAAACCCGTGGAGGACGGGCAGACCCGCGGTGCAAAATGTTTTTACAGC	15334						
Qy	585	GTGATCGGACAGATGAAGATGCTTCGACACGCTGCAGAAACACGACGT	631						
Db	15335	GTGATCGGACAGATGAAGATGCTTCGACACGCTGCAGAAACACGACGT	15381						

RESULT 2

RES001 2
US-09-453-702B-251/c
Sequence 251, Application US/09453702B
Patent No. 6365723
GENERAL INFORMATION:
APPLICANT: Blattner, Frederick R.
Burland, Valerie
Perna, Nicole T.
Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch. 1.44MB storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: word perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955

```

; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 251:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22306
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 251:
US-09-453-702B-251

      Query Match      79.4%;   Score 561.4;   DB 3;   Length 22306;
      Best Local Similarity 97.3%;   Pred. No. 1.7e-174;
      Matches 571;   Conservative 0;   Mismatches 16;   Indels 0;   Gaps 0;

QY      45  ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCAATCGCCCAAGC 104
      |||
DB      17193  ATCCAGTTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCAATCGCCCAAGC 17134
      |||

QY      105  TGGCGGTATCTGGGCATCGGGAGAGAAAGCCGTGCTTTTCCCGCGAGGTTGAAGCG 164
      |||
DB      17133  TGGCGGTATCTGGGCATCGGGAGAGAAAGCCGTGCTTTTCCCGCGAGGTTGAAGCG 17074
      |||

QY      165  GCATGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGAGCTTTGAGCGGAAAAAGCGACG 224
      |||
DB      17073  GCATGAAAGAGTTTGCCGAAATGACTGTTGCTGCTCATTTGAGCGGAAAAAGCGACG 17014
      |||

QY      225  TTTTACATGATGATTCGGGAAGGTGTGGCATGCAAGCTTTTAACGGTGAATGTTCTGTT 284
      |||
DB      17013  TTTTACATGATGATTCGGGAAGGTGTGGCATGCAAGCTTTTAACGGTGAATGTTCTGTT 16954
      |||

QY      285  CAGGCCACTCGGATACCAAGTTTCGTTCGGGCTTTTTCGGGACACAGTTCGGGATGGTCAAG 344
      |||
DB      16953  CAGGCCACTCGGATACCCGTCCTTCGGGACATGTTTCGGGACACAGTTCGGGATGGTCAAG 16894
      |||

QY      345  CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAATCGCGTCCGGGTGTG 404
      |||
DB      16893  CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAATCGCGTCCGGGTGTG 16834
      |||

QY      405  CAGATTAAATGACAGCGGTGCGCGCTGGGATATTACGTACAGGAGACGGGTATCCTCGC 464
      |||
DB      16833  CAGATTAAATGACAGCGGTGCGCGCTGGGATATTACGTACAGGAGACGGGTATCCTCGC 16774
      |||

QY      465  TGGATCCCGCAGAAATGGACATGGATACCCCGTAGTTTACCCGGCGGCGCGCTCGTTC 524
      |||
DB      16773  TGGATCCCGCAGAAATGGACATGGATACCCCGTAGTTTACCCGGCGGCGCGCTCGTTC 16714
      |||

QY      525  ATTCAGGTTTTGAACCGGTGAGAGACGGGACATCGCGGTGCAAAATGTTTTTACAGC 584
      |||
DB      16713  ATTCAGGTTTTGAACCGGTGAGAGACGGGACATCGCGGTGCAAAATGTTTTTACAGC 16654
      |||

QY      585  GTGATGGACGAGATGAAGATGCTGCACACGCTGCGAGAAACAGCGACGT 631
      |||
DB      16653  GTGATGGACGAGATGAAGATGCTGCACACGCTGCGAGAAACAGCGACGT 16607
      |||

RESULT 3
US-09-453-702B-96/c
; Sequence 96, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
;            Burland, Valerie
;            Perna, Nicole T.
;            Plunkett, Guy
;            Welch, Rod

```

TITLE OF INVENTION: No. 6365723el Sequences of E. coli 0157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:

ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI

COUNTRY: US
ZIP: 53701-2113

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/453,702B

FILING DATE: 03-Dec-1999

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/110,955

FILING DATE: 04-DEC-1998

ATTORNEY/AGENT INFORMATION:

NAME: Seay, Nicholas J.

REGISTRATION NUMBER: 27386

REFERENCE/DOCKET NUMBER: 960296.95017

TELECOMMUNICATION INFORMATION:

TELEPHONE: (608) 251-5000

TELEFAX: (608) 251-9166

INFORMATION FOR SEQ ID NO: 96:

SEQUENCE CHARACTERISTICS:

LENGTH: 34063

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 96:

US-09-453-702B-96

Query Match 31.5%; Score 223; DB 3; Length 34063;
Best Local Similarity 61.1%; Pred. No. 2.6e-62;
Matches 361; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

Qy 51 CTGATCAGGATCATATGCTGGGTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGGC 110
Db 14695 CTGCATAAGGATCATATGCTGGGTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGGC 14636

Qy 111 TATCTGGGCATCGGGAGGAGAACCGCGTCTTTTCCCGAGGTTTGAAGCGGCATGG 170
Db 14635 TGGCTGGGGATCGGGAGGAGAACCGCGTCTTTTCCCGAGGTTTGAAGCGGCATGG 14576

Qy 171 AAAGAGTTTCCGAGGATGACTGCTGCTGCAATTGACGTTGAGCGAAACGACGCTTTACC 230
Db 14575 TCGGAATACGCGAAGGATGCTGGCGAGATCGAGTGAAGGAAACGACGCTTACG 14516

Qy 231 ATGATGATTCGGGAGGTGCGCATGACACCCCTTTTAAAGTGAATGTTGTTTCAGGCC 290
Db 14515 GAATTTATTCGGTGAAGGTGCGCATGACACCCCTTTTAAAGTGAATGTTGTTTCAGGCC 14456

Qy 291 ACCTGGGATACAGTTCGTCGCGGCTTTCCGGACACAGTTCCGATGGTCAGCCCGAAG 350
Db 14455 GTCTGGGATACGGAACACCGAGTTATTCGTCAGCGTTTAAAGCGTGAGTCCGAAA 14396

Qy 351 CGCATCAGCAACCGAAACAATACCGCGGACAGCGGAACTCCGTCGCGGTGTCAGATT 410
Db 14395 CGGGTGGACACCGCAGGACACGGTATGGGAACCGTTTCTCGCGCGCGGGTGAGGTC 14336

Qy 411 AATGACCGGTGCGCGCTGGATATTAAGTACGAGGAGCGGTATCTCTGGTGGATG 470
Db 14335 GATCGATATGCGCGCTGGATATTAAGTACGAGGAGCGGTATCTCTGGTGGATG 14276

Qy 471 CCGCAGAAATCGATGATACCGGTGAGTTACCGCGCGCGCGCTGCTTCATTCAC 530
Db 14275 AGTGGATGAGGAAACCGATCCCGGTGAATCTCCACCGCGCGCTCCGGCCATCTGCAT 14216

Qy 531 GTTTTCAACCCGTGGAGGACGGCAGACTCGCGGTCCAAATGTGTTTTACGCGTGA 590
Db 14215 ATTTTCGAGCGGTGGAGGACGGCAGACTCGCGGTCCAAATGTGTTTTACGCGTGA 14156

Qy 591 GAGCAGATGAAGATGCTCGACACGCTCGAGAACACGCGAGCTAGATTAAACC 641
Db 14155 GAACGGCTGAAGATGCTCGATTCCTCGAGGCAACACAGCTTCAGTCGGCC 14105

RESULT 4

US-08-844-274-16/c

Sequence 16, Application US/08844274B

Patent No. 6218185

GENERAL INFORMATION:

APPLICANT: Fraser Jr., Malcom J.

APPLICANT: Shirk, Paul D.

APPLICANT: Elick, Teri A.

APPLICANT: Perera, Omathage

TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System

TITLE OF INVENTION: for Insects

FILE REFERENCE: 0148.96

CURRENT APPLICATION NUMBER: US/08/844,274B

CURRENT FILING DATE: 1997-04-18

EARLIER APPLICATION NUMBER: 60/016,234

EARLIER FILING DATE: 1996-04-19

NUMBER OF SEQ ID NOS: 22

SOFTWARE: PatentIn Ver. 2.0 - beta

SEQ ID NO 16

LENGTH: 5194

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: plasmid

OTHER INFORMATION: IFP2B/Xpuc18.1

US-08-844-274-16

Query Match 10.4%; Score 73.8; DB 3; Length 5194;

Best Local Similarity 83.2%; Pred. No. 2.3e-13;

Matches 84; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 605 GCTCGACACCGCTCGAGACACGCGAGTAGATTAAACCTAGAAAGATAATCATATTGTGAC 664
Db 2955 GCTCGTAGCCCGGGATCGGATCTCTAGATTAAACCTAGAAAGATAATCATATTGTGAC 2896

Qy 665 GTACGTTAAAGATAATCATCGTAAATAATTCAGCATGGAT 705
Db 2895 GTACGTTAAAGATAATCATCGTAAATAATTCAGCATGGAT 2855

RESULT 5

US-09-598-421-16/c

Sequence 16, Application US/09598421

Patent No. 6551825

GENERAL INFORMATION:

APPLICANT: Fraser Jr., Malcom J.

APPLICANT: Shirk, Paul D.

APPLICANT: Elick, Teri A.

APPLICANT: Perera, Omathage

TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System

TITLE OF INVENTION: for Insects

FILE REFERENCE: 0148.96

CURRENT APPLICATION NUMBER: US/09/598,421

CURRENT FILING DATE: 2000-06-19

PRIOR APPLICATION NUMBER: 60/016,234

PRIOR FILING DATE: 1996-04-19

NUMBER OF SEQ ID NOS: 22

SOFTWARE: PatentIn Ver. 2.0 - beta

SEQ ID NO 16

LENGTH: 5194

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

```
; OTHER INFORMATION: Description of Artificial Sequence:plasmid
; OTHER INFORMATION: IFF2B/Xpuc18.1
US-09-598-421-16

Query Match      10.4%; Score 73.8; DB 4; Length 5194;
Best Local Similarity 83.2%; Pred. No. 2.3e-13;
Matches 84; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 605 GCTCGACACGCTCAGAACACGAGCTAGATTAAACCTAGAAAGATAATCATATTGTGAC 664
Db      |||||
QY 2955 GCTCGGTACCCGGGGATCGGATCTCTAGATTAAACCTAGAAAGATAATCATATTGTGAC 2896
Db      |||||

QY 665 GTAGCTTAAAGATAATCATCGTAAATAATTGACGATGGAT 705
Db      |||||

QY 2895 GTAGCTTAAAGATAATCATCGTAAATAATTGACGATGTGT 2855
Db      |||||

RESULT 6
US-08-844-274-17/c
; Sequence 17, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omathhage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0 - beta
; SEQ ID NO 17
; LENGTH: 5194
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:plasmid
; OTHER INFORMATION: IFF2B/XsupF4H
US-08-844-274-17

Query Match      10.4%; Score 73.4; DB 3; Length 5194;
Best Local Similarity 87.9%; Pred. No. 3.1e-13;
Matches 80; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 615 CTGCAGAACACGAGCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTAA 674
Db      |||||
QY 4773 CTCTAGAGGGATCCTCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTAA 4714
Db      |||||

QY 675 GATAATCATCGTAAATAATTGACGATGGAT 705
Db      |||||

QY 4713 GATAATCATCGTAAATAATTGACGATGTGT 4683
Db      |||||

RESULT 7
US-09-598-421-17/c
; Sequence 17, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omathhage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19

; OTHER INFORMATION: Description of Artificial Sequence:plasmid
; OTHER INFORMATION: IFF2B/Xpuc18.1
US-09-598-421-16

Query Match      10.4%; Score 73.8; DB 4; Length 5194;
Best Local Similarity 83.2%; Pred. No. 2.3e-13;
Matches 84; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 605 GCTCGACACGCTCAGAACACGAGCTAGATTAAACCTAGAAAGATAATCATATTGTGAC 664
Db      |||||
QY 2955 GCTCGGTACCCGGGGATCGGATCTCTAGATTAAACCTAGAAAGATAATCATATTGTGAC 2896
Db      |||||

QY 665 GTAGCTTAAAGATAATCATCGTAAATAATTGACGATGGAT 705
Db      |||||

QY 2895 GTAGCTTAAAGATAATCATCGTAAATAATTGACGATGTGT 2855
Db      |||||

RESULT 6
US-08-844-274-17/c
; Sequence 17, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omathhage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0 - beta
; SEQ ID NO 17
; LENGTH: 5194
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:plasmid
; OTHER INFORMATION: IFF2B/XsupF4H
US-08-844-274-17

Query Match      10.4%; Score 73.4; DB 3; Length 5194;
Best Local Similarity 87.9%; Pred. No. 3.1e-13;
Matches 80; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 615 CTGCAGAACACGAGCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTAA 674
Db      |||||
QY 4773 CTCTAGAGGGATCCTCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTAA 4714
Db      |||||

QY 675 GATAATCATCGTAAATAATTGACGATGGAT 705
Db      |||||

QY 4713 GATAATCATCGTAAATAATTGACGATGTGT 4683
Db      |||||

RESULT 7
US-09-598-421-17/c
; Sequence 17, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omathhage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19

; OTHER INFORMATION: Description of Artificial Sequence:plasmid
; OTHER INFORMATION: IFF2B/XsupF4H
US-09-598-421-16

Query Match      10.4%; Score 73.4; DB 4; Length 5194;
Best Local Similarity 87.9%; Pred. No. 3.1e-13;
Matches 80; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 615 CTGCAGAACACGAGCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTAA 674
Db      |||||
QY 4773 CTCTAGAGGGATCCTCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTAA 4714
Db      |||||

QY 675 GATAATCATCGTAAATAATTGACGATGGAT 705
Db      |||||

QY 4713 GATAATCATCGTAAATAATTGACGATGTGT 4683
Db      |||||

RESULT 8
US-08-844-274-10
; Sequence 10, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omathhage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0 - beta
; SEQ ID NO 10
; LENGTH: 5679
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:p3E1.2 delta
; OTHER INFORMATION: TRL
US-08-844-274-10

Query Match      9.6%; Score 67.8; DB 3; Length 5679;
Best Local Similarity 97.2%; Pred. No. 2.3e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAAACCTAGAAAGATAATCATATTGTGACGTAAAGATAATCATCGTAA 694
Db      |||||
QY 1080 TTAAACCTAGAAAGATAATCATATTGTGACGTAAAGATAATCATCGTAA 1139
Db      |||||

QY 695 ACGATGGAT 705
Db      |||||

QY 1140 ACGCATGTGT 1150
Db      |||||

RESULT 9
US-09-598-421-10
; Sequence 10, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omathhage
```

```
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; CURRENT FILING DATE: 2000-06-19
; PRIOR FILING DATE: 60/016,234
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 10
; LENGTH: 5679
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:p3el.2 delta
; OTHER INFORMATION: TR
US-09-598-421-10

Query Match          9.6%; Score 67.8; DB 4; Length 5679;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 1080 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 1139

QY 695 ACGCATGGGAT 705
Db 1140 ACGCATGTGTT 1150

RESULT 10
US-08-844-274-13
; Sequence 13, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 13
; LENGTH: 6723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Clone p3el.2H/S
US-08-844-274-13

Query Match          9.6%; Score 67.8; DB 3; Length 6723;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 1080 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 1139

QY 695 ACGCATGGGAT 705
Db 1140 ACGCATGTGTT 1150

RESULT 11
US-08-844-274-14/c
; Sequence 14, Application US/08844274B
```

```
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 14
; LENGTH: 6723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:piggyBac/opd
US-08-844-274-14

Query Match          9.6%; Score 67.8; DB 3; Length 6723;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 5644 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 5585

QY 695 ACGCATGGGAT 705
Db 5584 ACGCATGTGTT 5574

RESULT 12
US-09-598-421-13
; Sequence 13, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 13
; LENGTH: 6723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Clone p3el.2H/S
US-09-598-421-13

Query Match          9.6%; Score 67.8; DB 4; Length 6723;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 1080 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 1139

QY 695 ACGCATGGGAT 705
Db 1140 ACGCATGTGTT 1150
```

4082 TTAACCCCTAGAAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 4023

RESULT 13
US-09-598-421-14/c
; Sequence 14, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; FILE REFERENCE: 0148.96
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 14
; LENGTH: 6723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:piggyBac/opd
US-09-598-421-14

Query Match 9.6%; Score 67.8; DB 4; Length 6723;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCCCTAGAAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
DB 5644 TTAACCCCTAGAAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 5585

QY 695 ACGCATGGGAT 705
DB 5584 ACGCATGTGTT 5574

RESULT 14
US-08-844-274-20/c
; Sequence 20, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 20
; LENGTH: 7560
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:p3E1.2hs/opd
US-08-844-274-20

Query Match 9.6%; Score 67.8; DB 3; Length 7560;
Best Local Similarity 97.2%; Pred. No. 2.7e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCCCTAGAAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694

Db 4082 TTAACCCCTAGAAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 4023
QY 695 ACGCATGGGAT 705
DB 4022 ACGCATGTGTT 4012

RESULT 15
US-09-598-421-20/c
; Sequence 20, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 20
; LENGTH: 7560
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:pjE1.2hs/opd
US-09-598-421-20

Query Match 9.6%; Score 67.8; DB 4; Length 7560;
Best Local Similarity 97.2%; Pred. No. 2.7e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCCCTAGAAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
DB 4082 TTAACCCCTAGAAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 4023

QY 695 ACGCATGGGAT 705
DB 4022 ACGCATGTGTT 4012

Search completed: June 5, 2005, 23:07:32
Job time : 122.806 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 13:38:04 ; Search time 2067.76 Seconds
(without alignments)
13014.775 Million cell updates/sec

Title: US-10-826-523-40

Perfect score: 707

Sequence: 1 ggaaccatgcgtcaatttt.....aaattgacgatggatcc 707

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	585.4	82.8	682	7	CK781302	CK781302 UI-M-GIO-
C 2	585.4	82.8	703	7	CF537771	CF537771 UI-M-GIO-
C 3	585.4	82.8	711	7	CF743678	CF743678 UI-M-GIO-
C 4	585.4	82.8	853	9	CR065566	CR065566 Forward s
C 5	584.4	82.7	703	5	BQ154655	BQ154655 NF087C061
C 6	584.4	82.7	719	6	CD350897	CD350897 UI-M-GIO-
C 7	583.8	82.6	751	9	CR087413	CR087413 Forward s
C 8	580.4	82.1	716	9	CR131675	CR131675 Forward s
C 9	580.4	82.0	786	9	CR108810	CR108810 Forward s
C 10	565.4	80.0	691	6	CD350776	CD350776 UI-M-GIO-
C 11	564.4	79.8	762	9	CR035207	CR035207 Reverse s
C 12	563.4	79.7	749	6	CD351273	CD351273 UI-M-GIO-
C 13	562.4	79.5	733	9	CR139473	CR139473 Forward s
C 14	555.4	78.5	633	9	CR167678	CR167678 Reverse s
C 15	554.4	78.4	791	6	CB520716	CB520716 UI-M-GIO-
C 16	550.4	77.9	630	9	CR013093	CR013093 Reverse s
C 17	549.4	77.7	702	9	CR084987	CR084987 Reverse s
C 18	545.4	77.1	611	4	B1423083	B1423083 EST533749
C 19	543.4	76.9	716	7	CK781284	CK781284 UI-M-GIO-
C 20	540.8	76.5	716	7	CF851544	CF851544 pMA007xC
C 21	537.2	76.0	706	1	AV731514	AV731514 AV731514
C 22	535.4	75.7	730	9	CR077673	CR077673 Reverse s
C 23	525.4	74.3	635	9	BX982794	BX982794 Forward s
C 24	521.8	73.8	536	9	CR026633	CR026633 Reverse s

ALIGNMENTS

RESULT 1
CK781302/c
LOCUS CK781302 682 bp mRNA linear EST 23-FEB-2004
DEFINITION UI-M-GIO-clp-g-05-0-UI.r1 NIH_BMAP_GIO Mus musculus CDNA clone
IMAGE:30619060 5', mRNA sequence.
ACCESSION CK781302
VERSION CK781302.1 GI:42746980
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 682)
AUTHORS NIH-MGC
TITLE NIH-MGC National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES

Seq primer: pyX-5.
Location/Qualifiers
1..682
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30619060"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpcc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GIO"
/note="Organ: Brain; Vector: pyX- Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pyX-Asc vector. The library tag

C 25 516 73.0 887 9 CR192748
C 26 515 72.8 885 9 CR255010
C 27 498.4 70.5 683 6 CD348896
C 28 493 69.7 495 4 BG457760
C 29 490.4 69.4 786 6 CB520098
C 30 482.6 68.3 795 6 CD351738
C 31 477.4 67.5 760 7 CF538326
C 32 474.4 67.1 620 9 CR034585
C 33 474.4 67.1 620 9 CR086282
C 34 474.4 67.1 620 9 CR171547
C 35 471.8 66.7 620 9 CR078862
C 36 467.4 66.1 752 9 CR088896
C 37 459.6 65.0 805 6 CD352530
C 38 457 64.6 676 5 BQ154673
C 39 455.4 64.4 483 9 CR068828
C 40 447.4 63.3 547 9 CR148906
C 41 447.4 63.3 885 9 CR024214
C 42 444.4 62.9 599 9 CR079016
C 43 443.4 62.7 702 9 CR038019
C 44 440.8 62.3 538 9 CR029330
C 45 437.6 61.9 901 6 CA328434

CR192748 Forward s
CR255010 Forward s
CD348896 UI-M-FY0-
BG457760 NF036R03P
CB520098 UI-M-GIO-
CD351738 UI-M-GIO-
CF538326 UI-M-GIO-
CR034585 Forward s
CR086282 Reverse s
CR171547 Reverse s
CR078862 Forward s
CR088896 Forward s
CD352530 UI-M-GIO-
BQ154673 NF095C111
CR068828 Reverse s
CR148906 Forward s
CR024214 Reverse s
CR079016 Forward s
CR038019 Forward s
CR029330 Forward s
CA328434 UI-M-FY0-

sequence located between the Not I site and the polyA tail is AGGACACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN		Query Match	82.8%;	Score 585.4;	DB 7;	Length 682;
		Best Local Similarity	99.8%;	Pred. No. 2.4e-164;		
		Matches 586;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	45	ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTTCATCGCCCAAGC	104			
Db	611	ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTTCATCGCCCAAGC	552			
QY	105	TGCGCTATCTGGGCATTCGGGAGGAAGACCCCGTGCCTTTTCCCGGAGGTTGAAGCG	164			
Db	551	TGCGCTATCTGGGCATTCGGGAGGAAGACCCCGTGCCTTTTCCCGGAGGTTGAAGCG	492			
QY	165	GCATGAAAGAGTTTCCCGAGGATGACTGCTGCTGATTGACGTTGAGCGAAGACGACG	224			
Db	491	GCATGAAAGAGTTTCCCGAGGATGACTGCTGCTGATTGACGTTGAGCGAAGACGACG	432			
QY	225	TTTACCATGATGATTCGGGAAGGTGTGGCCATGACGCGCTTTAAACGGTGAACCTGTT	284			
Db	431	TTTACCATGATGATTCGGGAAGGTGTGGCCATGACGCGCTTTAAACGGTGAACCTGTT	372			
QY	285	CAGGCCACCTGGGATACCGATGTCGCGGCTTTTCCGACACAGTTCGGATGTCGACG	344			
Db	371	CAGGCCACCTGGGATACCGATGTCGCGGCTTTTCCGACACAGTTCGGATGTCGACG	312			
QY	345	CCGAGCGCATCAGCAACCCGAAACATACCGCGACAGCGGAACCTGCCGTGCCGTG	404			
Db	311	CCGAGCGCATCAGCAACCCGAAACATACCGCGACAGCGGAACCTGCCGTGCCGTG	252			
QY	405	CAGATTAATGACAGCGGTGCGGCTGGGATATTACGTGAGGACGAGCGGTATCTCGC	464			
Db	251	CAGATTAATGACAGCGGTGCGGCTGGGATATTACGTGAGGACGAGCGGTATCTCGC	192			
QY	465	TGATGCCGAGAAATGGACATGATACCGCTGAGTTACCGCGGGCGCGCTCGTTC	524			
Db	191	TGATGCCGAGAAATGGACATGATACCGCTGAGTTACCGCGGGCGCGCTCGTTC	132			
QY	525	ATTACGTTTTTGAACCGTGGAGGACGGCAGACTCGCGTGAATGTGTTTTACAG	584			
Db	131	ATTACGTTTTTGAACCGTGGAGGACGGCAGACTCGCGTGAATGTGTTTTACAG	72			
QY	585	GTGATGGAGCAGATGAAGATGCTCGACAGCTGCAGAACACGACGCT	631			
Db	71	GTGATGGAGCAGATGAAGATGCTCGACAGCTGCAGAACACGACGCT	25			

RESULT 2
CF537771/c
LOCUS
DEFINITION
IMAGE:30537799 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CF537771 703 bp mRNA linear EST 12-SEP-2003
UI-M-GIO-chn-m-08-0-UI-r1 NIH_BMAP_GIO Mus musculus cDNA clone
IMAGE:30537799 5', mRNA sequence.
CF537771
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 703)
NIH-MGC <http://imgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.
Location/Qualifiers
1..703
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30537799"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (11 phage resistant)"
/clone_lib="NIH_BMAP_GIO"
/note="Organ: Brain; Vector: pYX-Asc; Site:1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGGACACAG. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Query Match	82.8%;	Score 585.4;	DB 7;	Length 703;
Best Local Similarity	99.8%;	Pred. No. 2.4e-164;		
Matches 586;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	45	ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTTCATCGCCCAAGC	104	
Db	643	ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTTCATCGCCCAAGC	584	
QY	105	TGCGCTATCTGGGCATTCGGGAGGAAGAAGCCCGTGCCTTTTCCCGGAGGTTGAAGCG	164	
Db	583	TGCGCTATCTGGGCATTCGGGAGGAAGAAGCCCGTGCCTTTTCCCGGAGGTTGAAGCG	524	
QY	165	GCATGGAAGAGTTTTCGGAGGATGACTGCTGCTGCATTGACGTTGACGGAAGACGCACG	224	
Db	523	GCATGGAAGAGTTTTCGGAGGATGACTGCTGCTGCATTGACGTTGACGGAAGACGCACG	464	
QY	225	TTTACCATGATGATTCGGGAAGGTGTGGCCATGACGCGCTTTAAACGGTGAACCTGTCGTT	284	
Db	463	TTTACCATGATGATTCGGGAAGGTGTGGCCATGACGCGCTTTAAACGGTGAACCTGTCGTT	404	
QY	285	CAGGCCACCTGGGATACCAAGTTCTGTCGCGGCTTTTCCGGAACACAGTTCGGATGTCGACG	344	
Db	403	CAGGCCACCTGGGATACCAAGTTCTGTCGCGGCTTTTCCGGAACACAGTTCGGATGTCGACG	344	
QY	345	CCGAAGCGCATCAGCAACCCGAAACAATAACCGGCGACAGCCGGAACTGCCGTGCCGTGTG	404	
Db	343	CCGAAGCGCATCAGCAACCCGAAACAATAACCGGCGACAGCCGGAACTGCCGTGCCGTGTG	284	
QY	405	CAGATTAATGACAGCGGTGCGGCTGGGATATTACGTGACGGAGGAACGGGTATCTCTGGC	464	
Db	283	CAGATTAATGACAGCGGTGCGGCTGGGATATTACGTGACGGAGGAACGGGTATCTCTGGC	224	
QY	465	TGGATGCCGAGAAATGGACATGATGATACCCCGTGAAGTTACCCGCGGCGCGCTCGTTC	524	
Db	223	TGGATGCCGAGAAATGGACATGATGATACCCCGTGAAGTTACCCGCGGCGCGCTCGTTC	164	
QY	525	ATTACGTTTTTGAACCCGTGGAGGACGGCAGACTCGCGGTGCAAAATGTGTTTTACAGC	584	

Db	163	ATTCA	CGTTTTT	TGAAC	CGTGG	AGGAC	GGGCA	GAAGT	CGCTG	CAAAAT	GTGTTT	TACAGC	104
Qy	585	GTGAT	GAGCA	GAGAT	GCTCG	ACAC	CGTGC	AGAA	CA	CGCAGCT	631		
Db	103	GTGAT	GAGCA	GATGA	GATGCT	CGAC	CGTGC	AGAA	CA	CGCAGCT	57		

RESULT 3	
CF743678/c	
LOCUS	711 bp mRNA linear EST 10-OCT-2003
DEFINITION	UI-M-GIO-clq-e-19-o-UI.r1 NIH BMAP_G10 Mus musculus CDNA clone IMAGE:30616722 5', mRNA sequence.
ACCESSION	CF743678
VERSION	CF743678.1 GI:37640017
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 711)
AUTHORS	NTH-GSC http://mgc.nci.nih.gov/.
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-x@mail.nih.gov Tissue Procurement: Dr. James Lin, University of Iowa cdNA Library preparation: Dr. M. Bento Soares, University of Iowa cdNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/mousefl.html This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

ORIGIN	program coordinator."
Query Match	82.8%; Score 585.4; DB 7; Length 711;
Best Local Similarity	99.8%; Pred. No. 2.4e-164;
Matches 586; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
Qy	45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCCTTTTTTCGGGTCAGTCAATGCCCAAGC 104
Db	654 ATCCAGCTGCATCAGGATCATATCGTCGGGTCCTTTTTTCGGGTCAGTCAATGCCCAAGC 595
Qy	105 TGGCGCTATCTGGGCATCGGGGAGGAAGACCCGTCCTTTTTCCCGCAGGTTGAAGCG 164
Db	594 TGGCGCTATCTGGGCATCGGGGAGGAAGACCCGTCCTTTTTCCCGCAGGTTGAAGCG 535

165	Qy	GCATGGAAGAGTTTGGCGAGGATGACTGCTGTCATTGACGTTGACGGAAACCGCAGC	224
534	Db	GCATGGAAGAGTTTGGCGAGGATGACTGCTGTCATTGACGTTGACGGAAACCGCAGC	475
225	Qy	TTTACCATGATGATTCCGGGAAGGTGTGGCCATGACAGCGCTTTTAAACGGTGAACTGTTTCGTT	284
474	Db	TTTACCATGATGATTCCGGGAAGGTGTGGCCATGACAGCGCTTTTAAACGGTGAACTGTTTCGTT	415
285	Qy	CAGGCCACTCGGGATACAGATTTCGTGCGGGCTTTTTCGGGACACAGTTTCCGGATGTCAGC	344
414	Db	CAGGCCACTCGGGATACAGATTTCGTGCGGGCTTTTTCGGGACACAGTTTCCGGATGTCAGC	355
345	Qy	CCGAAGCGCATCAGCAACCCGGAACAATACCGCGGACAGCCGGAACTGCGCGTGCCGGTGTG	404
354	Db	CCGAAGCGCATCAGCAACCCGGAACAATACCGCGGACAGCCGGAACTGCGCGTGCCGGTGTG	295
405	Qy	CAGATTAAATGACAGCGGTGCGGCGCTGGGATATTAACGTACGCGAGGACGGGTATCCTGGC	464
294	Db	CAGATTAAATGACAGCGGTGCGGCGCTGGGATATTAACGTACGCGAGGACGGGTATCCTGGC	235
465	Qy	TGGATGCCGAGAAATGGAATGATACCCGTGAGTTTACCCGCGGGCGGCCCTCGTTC	524
234	Db	TGGATGCCGAGAAATGGAATGATACCCGTGAGTTTACCCGCGGGCGGCCCTCGTTC	175
525	Qy	ATTACACGTTTTTGAAACCCGTGGAGGACGGGCAGAGACTCGCGGTGCAAAATGTTTTTACAGC	584
174	Db	ATTACACGTTTTTGAAACCCGTGGAGGACGGGCAGAGACTCGCGGTGCAAAATGTTTTTACAGC	115
585	Qy	GTGATGAGCAGATGAAGATGCTCGACACAGCTGCAGAACACGCGAGCT	631
114	Db	GTGATGAGCAGATGAAGATGCTCGACACAGCTGCAGAACACGCGAGCT	68

RESULT 4
 LOCUS CR065566
 DEFINITION Chromosome engineering clone MHP250k12, genomic survey sequence.
 ACCESSION CR065566
 VERSION 1
 KEYWORDS GSS; genome survey sequence; MICER.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 853)
 AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
 Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
 Rogers,J. and Bradley,A.
 TITLE Direct Submission
 JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. <http://www.sanger.ac.uk/MICER>
 FEATURES
 source
 1..853
 location/Qualifiers
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="MHP250k12"
 /clone_lib="MHP"

	Query Match	82.8%	Score 585.4	DB 9	Length 853
	Best Local Similarity	99.8%	Prsd. No. 2.5e-164		
	Matches 586	Conservative 0	Mismatches 1	Indels 0	Gaps 0
Qy	45	ATCTAGCTGCATCAGGATCATATCGTCGGGCTCTTTTTCCGGCTCAGTCATCGCCCAAGC	104		
Db	264	ATCCAGCTGCATCAGGATCATATCGTCGGGCTCTTTTTCCGGCTCAGTCATCGCCCAAGC	323		
Qy	105	TGCGGCTATCTGGGCATTCGGGGAGGAAGACCGCTGCTCTTTTCCCGCAGGTTGAAGCG	164		
Db	324	TGCGGCTATCTGGGCATTCGGGGAGGAAGACCGCTGCTCTTTTCCCGCAGGTTGAAGCG	383		

QY	165	GCATGGAAGAGTGTTCGCGAGATGACTGCTGCTGCAATTGACGTGTGAGCGGAAACGCACG	224
Db	384	GCATGGAAGAGTGTTCGCGAGATGACTGCTGCTGCAATTGACGTGTGAGCGGAAACGCACG	443
QY	225	TTTACCATGATGATTTCCGGGAAGGTGTGCCATGCA CGCTTTTAACGGTGAACATGTTTCGTT	284
Db	444	TTTACCATGATGATTTCCGGGAAGGTGTGCCATGCA CGCTTTTAACGGTGAACATGTTTCGTT	503
QY	285	CAGGCCACCTGGGATACCAAGTTCGTCGGCGCTTTTTCGGACACAGTTCCGGATGGTCAGC	344
Db	504	CAGGCCACCTGGGATACCAAGTTCGTCGGCGCTTTTTCGGACACAGTTCCGGATGGTCAGC	563
QY	345	CGGAAGCGCATCAGCAAAACCCGAAACAATACCGCGCACAGCCGGAACTCCCGTCGCGGTGTG	404
Db	564	CGGAAGCGCATCAGCAAAACCCGAAACAATACCGCGCACAGCCGGAACTCCCGTCGCGGTGTG	623
QY	405	CAGATTAATGACAGCGGTTCGGCGCTGGGATATTACGTACGACGAGGACGGGTATCTTGGC	464
Db	624	CAGATTAATGACAGCGGTTCGGCGCTGGGATATTACGTACGACGAGGACGGGTATCTTGGC	683
QY	465	TGGATGCCGACAGAAATGGACATGGATACCCCGTGAAGTTACCCGGCGGCGCGCTCGTTTC	524
Db	684	TGGATGCCGACAGAAATGGACATGGATACCCCGTGAAGTTACCCGGCGGCGCGCTCGTTTC	743
QY	525	ATTACAGTTTTTGAACCGTGGAGACGGGCAGACTCGCGTGCAAATGTGTTTACAGC	584
Db	744	ATTACAGTTTTTGAACCGTGGAGACGGGCAGACTCGCGTGCAAATGTGTTTACAGC	803
QY	585	GTGATGGACGATGAAGATGCTGCACACGCTGCAGAACACGCAGCT	631
Db	804	GTGATGGACGATGAAGATGCTGCACACGCTGCAGAACACGCAGCT	850

RESULT 5				
BQ154655/c				
LOCUS	BQ154655	703 bp	mRNA	linear EST 24-APR-2002
DEFINITION	NF087C06IR1F1050	Irradiated Medicago truncatula cDNA clone		
	NF087C06IR 5',	mRNA sequence.		

KEYWORDS	SOURCE	ORGANISM
EST.	Medicago truncatula (barrel medic)	Medicago truncatula
		Medicago truncatula
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE	1 (Pages 1 to 703)
AUTHORS	Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
TITLE	Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula irradiated library
JOURNAL	Unpublished (2001)
COMMENT	Contact: May GD Plant Biology Division The Samuel Roberts Noble Foundation 2510 Sam Noble Parkway, Ardmore, OK 73402, USA

```

FEATURES
source
1. 703
Location/Qualifiers
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF087C061R"
/tissue_type="seedlings"
/dev_stage="seedling"
/clone_lib="Irradiated"

```

/note=Vector: Lambda Zap; Seedlings were exposed either to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m² UV irradiation. Gamma-irradiated samples were harvested at 6, 12, 24 and 48 hours after treatment. UV-irradiated samples were harvested 24 hours post-treatment. cDNA was prepared from polyA⁺ enriched, pooled samples of equivalent amounts of total RNA from each sample. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using ExAssist helper phage and the E. coli strain XL1-blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."

ORIGIN

Query Match 82.7%; Score 584.4; DB 5; Length 703;
Best Local Similarity 99.7%; Pred. No. 4.8e-164;
Matches 585: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCCTTTTTTCGGGTCAGTCATCGCCCAAGC 104
|||||
Db 624 ATCCAGCTGCATCAGGATCATATCGTCGGGTCCTTTTTTCGGGTCAGTCATCGCCCAAGC 565

[illegible]

165	QY	GCATGAAAGAGATTTGCCGAGATGACTGCTGCTGCAATTCACCTTGAGCGAAACGCACG	224
504	Db	GCATGAAAGAGATTTGCCGAGATGACTGCTGCTGCAATTCACCTTGAGCGAAACGCACG	445

285	CAGGCCACCTGGGATACCACTTTCGTTCGCGCGCTTTTCCGGACACAGTTTCCGGATGGTCAGC	344
225	TTTACCATGATGATTCGGGAAAGGTGTGGCCATGTCACGCCCTTTAACGGTGAACATGTTCTGTT	284
444	TTTTCACATGATGATTCGGGAAAGGTGTGGCCATGTCACGCCCTTTAACGGTGAACATGTTCTGTT	385

[illegible]

Db	324	CCGAGCGGATCAGCAACCCGAA	CAATA	CGGGCAGACCGGAACTGCCGTGCGGGTGTG	265
Oy	405	CAGATTAA	CAGCGGTGCGCGCTGGGATATTAGT	CAGCGAGCAGCGGTATCCTGGC	464

Db	264	CAGATTAA	TGACAGCGGT	CGCGCTGGG	ATATTCGT	CAGCGAGG	ACGGGTATC	CTGGC	205
Qy	465	TGGATCCG	CAGAAATG	GACATGG	ATACCCCGG	CGGCGCGC	CTCGTTC	524	

Db	204	TGGATGCCGAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGGCTCGTTC	145
Qy	525	ATTACGTTTTTGAACCCCGTGGAGGACGGCAGACTCGCGGTGCAAAATGTTGTTTACAGC	584

Db
144 ATTACGTTTTGAACCGGTGAGGACGGGACACTCGCGGTCAATGTTTACAGC 85

QY
585 GTGATGGACGAGTGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631

Db 84 GTGATGGACAGATGAAGATGCTCGACAGCTGCAGAACACGCAGCT 38

RESULT 6
C350897 / 7

DEFINITION	UI-M-G10-cch-a-18-0-UI.r1 NIH-BWAP_G10 Mus musculus cDNA clone IMAGE:6853243 5', mRNA sequence.
ACCESSION	CD350897
VERSION	CD350897.1
KEYWORDS	GI:31142412
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
	Mus musculus
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 719)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1. .719
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6853243"
/tissue_type="whole brain"
/dev_stage="embryo 13.5, 14.5, 16.5, 17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_G10"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;
Site 2: Not I; The library was constructed according
Bonaïdo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with Ecor I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Query Match 82.7%; Score 584.4; DB 6; Length 719;
Best Local Similarity 99.7%; Pred. No. 4.8e-164;
Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 104
DB 663 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 604
QY 105 TGGCGCTATCTGGGCATCGGGAGGAAGAGCCGTCCTTTTTCGGCGAGGTTGAAGCG 164
DB 603 TGGCGCTATCTGGGCATCGGAGGAGAGAGCCGTCCTTTTTCGGCGAGGTTGAAGCG 544
QY 165 GCATGAAAGAGTTTCCGAGGATGACTGCTGCTGATTCGATTCGAGCGAAACGACG 224
DB 543 GCATGAAAGAGTTTCCGAGGATGACTGCTGCTGATTCGATTCGAGCGAAACGACG 484
QY 225 TTATCATCATGATTCGGGAAGGTGGCCATGACGCTTTTAAACGTTGTCGTT 284
DB 483 TTATCATCATGATTCGGGAAGGTGGCCATGACGCTTTTAAACGTTGTCGTT 424
QY 285 CAGGCCACCTGGGATACCAAGTTTCGTCGGCTTTTTCGGACACAGTTCCGGATGGTCAGC 344
DB 423 CAGGCCACCTGGGATACCAAGTTTCGTCGGCTTTTTCGGACACAGTTCCGGATGGTCAGC 364
QY 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAATTCGCGTGCCTGTG 404
DB 363 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAATTCGCGTGCCTGTG 304
QY 405 CAGATTAAATGACAGCGGTTCGGCGCTGGGATATTAGTCAGCGAGGACGGGTATCTCTGC 464

Db 303 CAGATTAAATGACAGCGGTTCGGCGCTGGGATATTAGTCAGCGAGGACGGGTATCTCTGC 244
QY 465 TGGATGCCCGAGAAATGGACATGATACCCCGTACGTTACCCGCGGGCGCTCGTTC 524
Db 243 TGGATGCCCGAGAAATGGACATGATACCCCGTACGTTACCCGCGGGCGCTCGTTC 184
QY 525 ATTACAGTTTGAACCCGTCGGAGGACGGGACGACTCGGTCGCAATGTTTACAGC 584
Db 183 ATTACAGTTTGAACCCGTCGGAGGACGGGACGACTCGGTCGCAATGTTTACAGC 124
QY 585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAAACGCGAGCT 631
Db 123 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAAACGCGAGCT 77
RESULT 7
CR087413/c
LOCUS CR087413 751 bp DNA linear GSS 05-JUL-2004
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and
chromosome engineering clone MHP438f24, genomic survey sequence.
ACCESSION CR087413
VERSION CR087413.1 GI:49821005
KEYWORDS GSS; Genome survey sequence; MICE.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 751)
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. <http://www.sanger.ac.uk/MICE>
FEATURES
1..751
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHP438f24"
/clone_lib="MHPp"

ORIGIN

Query Match 82.6%; Score 583.8; DB 9; Length 751;
Best Local Similarity 99.7%; Pred. No. 7.3e-164;
Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 104
DB 623 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 584
QY 105 TGGCGCTATCTGGGCATCGGGAGGAAGAGCCGTCCTTTTTCGGCGAGGTTGAAGCG 164
DB 563 TGGCGCTATCTGGGCATCGGGAGGAGAGAGCCGTCCTTTTTCGGCGAGGTTGAAGCG 504
QY 165 GCATGAAAGAGTTTCCGAGGATGACTGCTGCTGATTCGATTCGAGCGAAACGACG 224
DB 503 GCATGAAAGAGTTTCCGAGGATGACTGCTGCTGATTCGATTCGAGCGAAACGACG 444
QY 225 TTATCATCATGATTCGGGAAGGTGGCCATGACGCTTTTAAACGTTGTCGTT 284
DB 443 TTATCATCATGATTCGGGAAGGTGGCCATGACGCTTTTAAACGTTGTCGTT 384
QY 285 CAGGCCACCTGGGATACCAAGTTTCGTCGGCTTTTTCGGACACAGTTCCGGATGGTCAGC 344
DB 383 CAGGCCACCTGGGATACCAAGTTTCGTCGGCTTTTTCGGACACAGTTCCGGATGGTCAGC 324
QY 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAATTCGCGTGCCTGTG 404
DB 323 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAATTCGCGTGCCTGTG 264
QY 405 CAGATTAAATGACAGCGGTTCGGCGCTGGGATATTAGTCAGCGAGGACGGGTATCTCTGC 464

```
Db      263 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 204
Qy      465 TGGATGCCCGCAGAAATGGACATGGATACCCCGTGGAGTTACCCGGGGCGCGCCCTCGTTTC 524
Db      203 TGGATGCCCGCAGAAATGGACATGGATACCCCGTGGAGTTACCCGGGGCGCGCCCTCGTTTC 144
Qy      525 ATTACAGTTTTTTGAACCCCGTGGAGGACGGCGAGACTCGCGGTGCAAAATGTGTTTACAGC 584
Db      143 ATTACAGTTTTTTGAACCCCGTGGAGGACGGCGAGACTCGCGGTGCAAAATGTGTTTACAGC 84
Qy      585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGAGCT 631
Db      83 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGAGCT 37

RESULT 8
CR131675/c
LOCUS   CR131675               716 bp      DNA      linear      GSS 06-JUL-2004
DEFINITION
Forward strand read from insert in 3'HPRT insertion targeting and
chromosome engineering clone MHP438f23, genomic survey sequence.
ACCESSION
CR131675.1 GI:49879128
VERSION
GSS; genome survey sequence; MICER.
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 716)
AUTHORS
Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
Direct Submission
JOURNAL
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES
Location/Qualifiers
source
1..716
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHP438f23"
/clone_lib="MHPP"

ORIGIN
Query Match 82.1%; Score 580.4; DB 9; Length 716;
Best Local Similarity 99.8%; Pred. No. 7.6e-163;
Matches 591; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGC 104
Db      582 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGC 523
Qy      105 TGGCGCTATCTGGGCATCGGGAGGAGAACCGCGTCCCTTTTCCGCGAGGTTGAAGCG 164
Db      522 TGGCGCTATCTGGGCATCGGGAGGAGAACCGCGTCCCTTTTCCGCGAGGTTGAAGCG 463
Qy      165 GCATGAAAGAGTTTCCCGCAGGATCACTGCTGTCATTGAGCTTTGAGCGAAAACGCACG 224
Db      462 GCATGAAAGAGTTTCCCGCAGGATGACTGCTGCTGCAATTGAGCTTGAGCGAAAACGCACG 403
Qy      225 TTATCCATGATGATTGGGAAAGGTGTGGCCATGACACGCTTTTAAACGGTGAACTGTTCGTT 284
Db      402 TTATCCATGATGATTGGGAAAGGTGTGGCCATGACACGCTTTTAAACGGTGAACTGTTCGTT 343
Qy      285 CAGGCCACTCTGGATACAGTTTGTTCGGCGCTTTTCCGACACAGTTCCGGATGTCAGC 344
Db      342 CAGGCCACTCTGGATACAGTTTGTTCGGCGCTTTTCCGACACAGTTCCGGATGTCAGC 283
Qy      345 CCGAAGCGCATCAGCAACCCGAAACAATACCGCGCAGCAGCGGAACCTGCCGTGCCCGGTG 404
Db      282 CCGAAGCGCATCAGCAACCCGAAACAATACCGCGCAGCAGCGGAACCTGCCGTGCCCGGTG 223
Qy      405 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
```

```
Db      222 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 163
Qy      465 TGGATGCCCGCAGAAATGGACATGGATACCCCGTGGAGTTACCCGGGGCGCGCCCTCGTTTC 524
Db      162 TGGATGCCCGCAGAAATGGACATGGATACCCCGTGGAGTTACCCGGGGCGCGCCCTCGTTTC 103
Qy      525 ATTACAGTTTTTTGAACCCCGTGGAGGACGGCGAGACTCGCGGTGCAAAATGTGTTTACAGC 584
Db      102 ATTACAGTTTTTTGAACCCCGTGGAGGACGGCGAGACTCGCGGTGCAAAATGTGTTTACAGC 43
Qy      585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACG 626
Db      42 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACG 1

RESULT 9
CR108810/c
LOCUS   CR108810               786 bp      DNA      linear      GSS 05-JUL-2004
DEFINITION
Forward strand read from insert in 3'HPRT insertion targeting and
chromosome engineering clone MHP22016, genomic survey sequence.
ACCESSION
CR108810
VERSION
CR108810.1 GI:49856225
KEYWORDS
GSS; genome survey sequence; MICER.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 786)
AUTHORS
Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
Direct Submission
JOURNAL
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES
Location/Qualifiers
source
1..786
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHP22016"
/clone_lib="MHPP"

ORIGIN
Query Match 82.0%; Score 580; DB 9; Length 786;
Best Local Similarity 100.0%; Pred. No. 1e-162;
Matches 580; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      52 TCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGCTGGCGCT 111
Db      786 TGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGCTGGCGCT 727
Qy      112 ATCTGGGCATCGGGAGGAGAACCGTGCCTTTTCCGCGAGGTTGAAGCGCATGA 171
Db      726 ATCTGGGCATCGGGAGGAGAACCGTGCCTTTTCCGCGAGGTTGAAGCGCATGA 667
Qy      172 AAGAGTTTCCCGCAGGATGACTGCTGCTGCAATTGAGCTTTGAGCGAAAACGCAGTTTACCA 231
Db      666 AAGAGTTTCCCGCAGGATGACTGCTGCTGCAATTGAGCTTTGAGCGAAAACGCAGTTTACCA 607
Qy      232 TGATGATTCGGGAAAGGTGTGGCCATGCACGCTTTTAAACGGTGAACTGTTCGTTCAGGCCA 291
Db      606 TGATGATTCGGGAAAGGTGTGGCCATGCACGCTTTTAAACGGTGAACTGTTCGTTCAGGCCA 547
Qy      292 CCTGGGATACCAAGTTCTGTCGGGCTTTTCCGACACAGTTCCGGATGTCAGCCGGAAGC 351
Db      546 CTTGGGATACCAAGTTCTGTCGGGCTTTTCCGACACAGTTCCGGATGTCAGCCGGAAGC 487
Qy      352 GCATCAGCAACCCGAAACAATACCGCGCAGCAGCGGAACCTGCCGTGCCGTGTGCAGATTA 411
Db      486 GCATCAGCAACCCGAAACAATACCGCGCAGCAGCGGAACCTGCCGTGCCGTGTGCAGATTA 427
Qy      412 ATGACAGCGGTGCGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGATGC 471
```

Db	426	ATGACACGGGTCGGCGCTGGGATATTACGTACGGAGGACGGGTATCCTCGCTGGATGC	367
Qy	472	CGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCCTCGTTTCATTCA	531
Db	366	CGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCCTCGTTTCATTCA	307
Qy	532	TTTTTTGAACCCGTGGAGGACGGGCGAGACTCGCGGTGCAAAATGTGTTTTACACGCTGA	591
Db	306	TTTTTTGAACCCGTGGAGGACGGGCGAGACTCGCGGTGCAAAATGTGTTTTACACGCTGA	247
Qy	592	AGCAGATGAAGATGCTCGACACGCTGCAGAAACACGCAGCT	631
Db	246	AGCAGATGAAGATGCTCGACACGCTGCAGAAACACGCAGCT	207

RESULT 10
 CD350776/c
 LOCUS
 DEFINITION
 CD350776 691 bp mRNA linear EST 09-JUL-2003
 UI-M-G10-cgh-g-19-0-UI.r1 NIH BMAP_G10 Mus musculus cDNA clone
 IMAGE:6853388 5', mRNA sequence.
 ACCESSION
 CD350776
 VERSION
 CD350776.1 GI:31142363
 KEYWORDS
 EST.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 691)
 REFERENCE
 NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cqapbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousef1.html>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

FEATURES
source

ORIGIN

Query Match 80.0%; Score 565.4; DB 6; Length 691;
Best Local Similarity 99.8%; Pred. No. 2.4e-158;
Matches 566: Conservative 0; Mismatches 1; Indels 0

Qy	45	ATCTAGCTGCATCAGGATCATATCGTGGGTCTTTTTTCCGGCTCAGTTCATCGCCCAAGC	104
Db	567	ATCCAGCTGCATCAGGATCATATCGTGGGTCTTTTTTCCGGCTCAGTTCATCGCCCAAGC	508
Qy	105	TGGCGCTATCTGGGCATCGGGGAGGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG	164
Db	507	TGGCGCTATCTGGGCATCGGGGAGGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG	448
Qy	165	GCATGSAAGAGTTTGGCCGAGGTGACTGCTGCTGCATTGACGTGTAGCGAAGAAACGCACG	224
Db	447	GCATGSAAGAGTTTGGCCGAGGTGACTGCTGCTGCATTGACGTGTAGCGAAGAAACGCACG	388
Qy	225	TTTACCATGATGATTCGGGAAGGTGTGGCCATGACACGCCCTTTAACGGTGAACCTGTCGT	284
Db	387	TTTACCATGATGATTCGGGAAGGTGTGGCCATGACACGCCCTTTAACGGTGAACCTGTCGT	328
Qy	285	CAGGCCACCTGGGATACACAGTTCTGCGGGCTTTTTCGGGACACAGTTCCGGATGTCAGC	344
Db	327	CAGGCCACCTGGGATACACAGTTCTGCGGGCTTTTTCGGGACACAGTTCCGGATGTCAGC	268
Qy	345	CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAATCGCGTGCCGGTGTG	404
Db	267	CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAATCGCGTGCCGGTGTG	208
Qy	405	CAGATTAAATGACACGGTGTGGGCGCTGGGATATTACGTACAGCAGGACGGGTATCCTGGC	464
Db	207	CAGATTAAATGACACGGTGTGGGCGCTGGGATATTACGTACAGCAGGACGGGTATCCTGGC	148
Qy	465	TGGATCGCGAGAAATGGACATGATACCCCGTCAGTTACCCGCGGGCGGCCCTCGTTC	524
Db	147	TGGATCGCGAGAAATGGACATGATACCCCGTCAGTTACCCGCGGGCGGCCCTCGTTC	88
Qy	525	ATTCAACGTTTTTGAACCCGTGGAGGACGGGACAGACTCGCGGTGCAAAATGTGTTTTACAGC	584
Db	87	ATTCAACGTTTTTGAACCCGTGGAGGACGGGACAGACTCGCGGTGCAAAATGTGTTTTACAGC	28
Qy	585	GTGATGAGCAGATGAAGATGCTCGAC	611
Db	27	GTGATGAGCAGATGAAGATGCTCGAC	1

RESULT 11	TITLE
CR035207	JOURNAL
LOCUS	FEATURES
DEFINITION	SOURCE
ACCESSION	ORIGIN
VERSION	
KEYWORDS	
SOURCE	
ORGANIS	
REFERENCE	
AUTHORS	

Query Match 79.8%; Score 564; DB 9; Length 762;
Best Local Similarity 100.0%; Pred. No. 6.5e-158;
Matches 564; Conservative 0; Mismatches 0 Indels

Query Match 80.0%; Score 565.4; DB 6; Length 691;
 Best Local Similarity 99.8%; Pred. No. 2.4e-158;
 Matches 566: Conservative 0; Mismatches 1; Indels 0; Gaps 0;


```
QY 68 CGTCGGGCTCTTTTCCCGGCTCAGTCATCGCCCAAGCTCGCGCTATCTCGGCATCGGGGA 127
Db 2 CGTCGGGCTCTTTTCCCGGCTCAGTCATCGCCCAAGCTCGCGCTATCTCGGCATCGGGGA 61
QY 128 GGAAGAAGCCCGTGCCTTTTCCCGGAGGTTGAAGCGGCATGGAAGAGATTTGCCGAGGA 187
Db 62 GGAAGAAGCCCGTGCCTTTTCCCGGAGGTTGAAGCGGCATGGAAGAGATTTGCCGAGGA 121
QY 188 TGACTGCTGCTGATGAGTTGAGGGAAGCGCAGTTTACCATGATGATTCGGGAAGG 247
Db 122 TGACTGCTGCTGATGAGTTGAGGGAAGCGCAGTTTACCATGATGATTCGGGAAGG 181
QY 248 TGTGGCCATGACGCTTTTAAAGCGTGAAGTGTTCAGGCGCACCTGGGATACAGTTTC 307
Db 182 TGTGGCCATGACGCTTTTAAAGCGTGAAGTGTTCAGGCGCACCTGGGATACAGTTTC 241
QY 308 GTCCGGGCTTTTCCGACACAGTTCCGGATGTGACCCGCGGCATGAGCAACCCGAA 367
Db 242 GTCCGGGCTTTTCCGACACAGTTCCGGATGTGACCCGCGGCATGAGCAACCCGAA 301
QY 368 CAATACCGCGCAGCGGCACTGCGGTGCGCGTGTGCGAGATTAATGACAGCGGTGCGGC 427
Db 302 CAATACCGCGCAGCGGCACTGCGGTGCGCGTGTGCGAGATTAATGACAGCGGTGCGGC 361
QY 428 GCTGGGATATTACGTGACGCGAGCGGATCTCTGCTGATGCCGACGAAATGACATG 487
Db 362 GCTGGGATATTACGTGACGCGAGCGGATCTCTGCTGATGCCGACGAAATGACATG 421
QY 488 GATACCCCGTGTGATACCCGCGCGCGCCCTCGTTCAATTCACGTTTGAACCCGTTGA 547
Db 422 GATACCCCGTGTGATACCCGCGCGCGCCCTCGTTCAATTCACGTTTGAACCCGTTGA 481
QY 548 GGAAGCGGACACTCGGCGTGAATGTGTTTACAGCGTGTGAGCAGATGAAGATGCT 607
Db 482 GGAAGCGGACACTCGGCGTGAATGTGTTTACAGCGTGTGAGCAGATGAAGATGCT 541
QY 608 CGACACGCTGCAGAACCCGACGCT 631
Db 542 CGACACGCTGCAGAACCCGACGCT 565
```

```
RESULT 12
CD351273/c
LOCUS
DEFINITION
  UI-M-G10-cgh-f-22-0-UI.r1 NIH_BMAP_G10 Mus musculus cDNA clone
  IMAGE:6853367 5', mRNA sequence.
ACCESSION
  CD351273.1 GI:31142608
VERSION
  EST.
KEYWORDS
  Mus musculus (house mouse)
SOURCE
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 749)
  NIH-MGC http://mgi.nci.nih.gov/
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-remail.nih.gov
  Tissue procurement: Dr. Jim Lin, University of Iowa
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Distribution information can be found at
  http://genome.uiowa.edu/distribution/mousefl.html
  This clone was contributed by the Brain Molecular Anatomy Project
  (BMAP)
```

```
Seq primer: pYX-5
Location/Qualifiers
  1..749
  /organism="Mus musculus"
  /mol_type="mRNA"
```

```
FEATURES
  source
```

```
RESULT 13
CR139473/c
LOCUS
DEFINITION
```

```
Forward strand read from insert in 3'HPRT insertion targeting and
CR139473 733 bp DNA linear GSS 06-JUL-2004
```

```
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6853367"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP G10"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is ACCGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."
```

ORIGIN

```
Query Match 79.7%; Score 563.4; DB 6; Length 749;
Best Local Similarity 99.8%; Pred. No. 9.8e-158;
Matches 564; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 45 ATCTAGCTGCATCAGATCATATCGTCGGGCTCTTTTCCGGCTCAGTCATCGCCCAAGC 104
Db 565 ATCCAGCTGCATCAGGATCATATCGTCGGGCTCTTTTCCGGCTCAGTCATCGCCCAAGC 506
QY 105 TGGCGCTATCTCGGCATCGGGGAGGAAGAGCCGTCCTTTTCCGCGAGGTTGAAGCG 164
Db 505 TGGCGCTATCTCGGCATCGGGGAGGAAGAGCCGTCCTTTTCCGCGAGGTTGAAGCG 446
QY 165 GCATGAAAGAGTTTCCGAGGATGACTGCTGCTCATTCGATTCGAGCGAAACGACG 224
Db 445 GCATGAAAGAGTTTCCGAGGATGACTGCTGCTCATTCGATTCGAGCGAAACGACG 386
QY 225 TTTACATGATGATTCGGGAGGTTGGCCATGACGCGCTTTTAAAGGTGAACGTTCTGTT 284
Db 385 TTTACATGATGATTCGGGAGGTTGGCCATGACGCGCTTTTAAAGGTGAACGTTCTGTT 326
QY 285 CAGGCCACTCTGGATACAGATTCGTCGGGCTCTTTTCCGACACAGTTCCGATGTCAGC 344
Db 325 CAGGCCACTCTGGATACAGATTCGTCGGGCTCTTTTCCGACACAGTTCCGATGTCAGC 266
QY 345 CCGAAGCGCATCAGCAACCCGAAACATACCGGCGACAGCCGGAACCTGCCGTGCCGTGTG 404
Db 265 CCGAAGCGCATCAGCAACCCGAAACATACCGGCGACAGCCGGAACCTGCCGTGCCGTGTG 206
QY 405 CAGATTATGACAGCGGTGCGGCTGGGATATTACGTCAGCGAGCGGATCTCTGCG 464
Db 205 CAGATTATGACAGCGGTGCGGCTGGGATATTACGTCAGCGAGCGGATCTCTGCG 146
QY 465 TGGATGCCGCAAAATGGACATGATATCCCGTGAGTTTACCCGCGGCGCGCTCTGTTTC 524
Db 145 TGGATGCCGCAAAATGGACATGATATCCCGTGAGTTTACCCGCGGCGCGCTCTGTTTC 86
QY 525 ATTCACGTTTTTGAACCCCGTGGAGACCGGCGAGACTCGCGGTGCAAAATGTGTTTACAGC 584
Db 85 ATTCACGTTTTTGAACCCCGTGGAGACCGGCGAGACTCGCGGTGCAAAATGTGTTTACAGC 26
QY 585 GTGATGAGCAGATGAAGATGCTCG 609
Db 25 GTGATGAGCAGATGAAGATGCTCG 1
```


chromosome engineering clone MHPP274b23, genomic survey sequence.

CR139473
VERSION GI:49887345
KEYWORDS GSS; genome survey sequence; MICR.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 733)
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.
Direct Submission
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICR

TITLE
JOURNAL
FEATURES
Location/Qualifiers

source
1. .733
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPP274b23"
/clone_lib="MHPP"

ORIGIN

Query Match 79.5%; Score 562; DB 9; Length 733;
Best Local Similarity 99.3%; Pred. No. 2.6e-157;
Matches 573; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 56 TCAGGATCATATCGTCGGGTCTTTTTCGGGTCAGTCATCGCCCAAGCTGGCGCTATCT 115
DB 733 TCAGGATCATATCGTCGGGTCTTTTTCGGGTCAGTCATCGCCCAAGCTGGCGCTATCT 674
QY 116 GGCATCGGGAGGAAGACCGCTGCTCTTTCCCGCGAGGTTGAAGCGGCATGGAAGA 175
DB 673 GGCATCGGGAGGAAGACCGCTGCTCTTTCCCGCGAGGTTGAAGCGGCATGGAAGA 614
QY 176 GTTTCGGAGGATGACTGCTGTCGATTCAGTTCAGGAGAAACGACAGCTTTACCATGAT 235
DB 613 GTTTCGGAGGATGACTGCTGTCGATTCAGTTCAGGAGAAACGACAGCTTTACCATGAT 554
QY 236 GATTCCGGAGGTTGGCGCATGACGCCCTTTTACCGTGAACTGTTTCAGGCCACCTG 295
DB 553 GATTCCGGAGGTTGGCGCATGACGCCCTTTTACCGTGAACTGTTTCAGGCCACCTG 494
QY 296 GGATACCACTGTCGCGGCTTTTCGGACACAGTTTCGGATGTCAGCCCGAAGCGCAT 355
DB 493 GGATACCACTGTCGCGGCTTTTCGGACACAGTTTCGGATGTCAGCCCGAAGCGCAT 434
QY 356 CAGCAACCCGACAAATACCGGACAGCCGGAACCTCCG - TGCCGGTGTGCAGATTAATG 414
DB 433 CAGCAACCCGACAAATACCGGACAGCCGGAACCTCCGNNCTGCGNCTGTCAGATTAATG 374
QY 415 ACAGCGGTGCGGCTGGGATATTACGTCAGGAGGACGGGTATCTGGCTGGATCCCGC 474
DB 373 ACAGCGGTGCGGCTGGGATATTACGTCAGGAGGACGGGTATCTGGCTGGATCCCGC 314
QY 475 AGAAATGACATGATACCCCGTGAAGTTACCCGGCGGGCGGCTCGTTCAITTCAGGTTT 534
DB 313 AGAAATGACATGATACCCCGTGAAGTTACCCGGCGGGCGGCTCGTTCAITTCAGGTTT 254
QY 535 TTGAACCCGTGAGGACCGGACAGCTCGCGGTGCAAAATGTTTACAGCGTGATGGAGC 594
DB 253 TTGAACCCGTGAGGACCGGACAGCTCGCGGTGCAAAATGTTTACAGCGTGATGGAGC 194
QY 595 AGATGAAGATGCTCGACAGCTGCAGAACACGAGCT 631
DB 193 AGATGAAGATGCTCGACAGCTGCAGAACACGAGCT 157

RESULT 14
CR167678/c
LOCUS 633 bp DNA linear GSS 06-JUL-2004
DEFINITION Reverse strand read from insert in 3'HPRT insertion targeting and

chromosome engineering clone MHPP161k02, genomic survey sequence.

CR167678
VERSION GI:49946527
KEYWORDS GSS; genome survey sequence; MICR.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 633)
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.
Direct Submission
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICR

TITLE
JOURNAL
FEATURES
Location/Qualifiers

source
1. .633
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPP161k02"
/clone_lib="MHPP"

ORIGIN

Query Match 78.5%; Score 555; DB 9; Length 633;
Best Local Similarity 100.0%; Pred. No. 3.2e-155;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 TTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGAGGAAGC 136
DB 633 TTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGAGGAAGC 574
QY 137 CCGTGCCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGGTTTTCGCGAGATGCTGCTG 196
DB 573 CCGTGCCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGGTTTTCGCGAGATGCTGCTG 514
QY 197 CTGCATTGAGCTTGAGCGAAACGACGCTTTTACCATGATTCGCGAGAGTGTGGCAT 256
DB 513 CTGCATTGAGCTTGAGCGAAACGACGCTTTTACCATGATTCGCGAGAGTGTGGCAT 454
QY 257 GCACGCGCTTTTAAACGGTGAACCTGTTCTCAGGCCACCTGGGATACCAATTCGTCGGGCT 316
DB 453 GCACGCGCTTTTAAACGGTGAACCTGTTCTCAGGCCACCTGGGATACCAATTCGTCGGGCT 394
QY 317 TTTCCGGACACAGTTTCCGGATGTCAGCCCGAAGCGCATCAGCAACCCGAAACAATACCGG 376
DB 393 TTTCCGGACACAGTTTCCGGATGTCAGCCCGAAGCGCATCAGCAACCCGAAACAATACCGG 334
QY 377 CGACAGCGGAACCTGCGGTGCGGATTAATGACAGCGGTGCGGCGCTGGGATA 436
DB 333 CGACAGCGGAACCTGCGGTGCGGATTAATGACAGCGGTGCGGCGCTGGGATA 274
QY 437 TTACGTCAGCGAGGACGGGTATCTCGGCTGGATCCGCGAGAAATGACATGATACCCCG 496
DB 273 TTACGTCAGCGAGGACGGGTATCTCGGCTGGATCCGCGAGAAATGACATGATACCCCG 214
QY 497 TGAGTTTACCCGGCGGCGGCTCGCTTCATTTCAGCTTTTGAACCCGTCGAGGACGGCA 556
DB 213 TGAGTTTACCCGGCGGCGGCTCGCTTCATTTCAGCTTTTGAACCCGTCGAGGACGGCA 154
QY 557 GACTCGCGTCAAAATGTTTTCAGCGCTGATGGAGCAGATGAAGATGCTCAGACGCT 616
DB 153 GACTCGCGTCAAAATGTTTTCAGCGCTGATGGAGCAGATGAAGATGCTCAGACGCT 94
QY 617 GCAGAACACGAGCT 631
DB 93 GCAGAACACGAGCT 79

RESULT 15
CR520716/c
LOCUS 791 bp mRNA linear EST 09-JUL-2003
DEFINITION UI-M-G10-cej-j-06-0-UI.r1 NIH_BMAP_G10 Mus musculus cDNA clone

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 19:22:55 ; Search time 1994.86 Seconds
(without alignments)
11284.911 Million cell updates/sec

Title: US-10-826-523-41

Perfect score: 3662

Sequence: 1 ctaaatgtaagcgtaata.....atttccccgaaagtgcac 3662

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 5706582 seqs, 3073711274 residues

Total number of hits satisfying chosen parameters: 11413164

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3662	100.0	3662	13	US-10-001-189-41
2	2665.4	72.8	3637	15	US-10-211-079-18
3	2665.4	72.8	3637	16	US-10-356-708-32
4	2665.4	72.8	3637	18	US-10-280-913A-32
5	2665.4	72.8	3637	18	US-10-684-134-32
6	2665.4	72.8	3637	18	US-10-637-758-32
7	2664.8	72.8	3637	13	US-10-066-390-4
8	2664.8	72.8	3637	13	US-10-206-030-4
9	2664.8	72.8	3637	15	US-10-211-079-4
10	2664.8	72.8	3637	16	US-10-356-708-17
11	2664.8	72.8	3637	16	US-10-205-772-4
					Sequence 41, Appl
					Sequence 18, Appl
					Sequence 32, Appl
					Sequence 32, Appl
					Sequence 32, Appl
					Sequence 4, Appl
					Sequence 4, Appl
					Sequence 17, Appl
					Sequence 4, Appl

C 12	2664.8	72.8	3637	18	US-10-280-913A-17	Sequence 17, Appl
C 13	2664.8	72.8	3637	18	US-10-684-134-17	Sequence 17, Appl
C 14	2664.8	72.8	3637	18	US-10-637-758-17	Sequence 17, Appl
C 15	2656.4	72.5	3637	13	US-10-066-390-3	Sequence 3, Appl
C 16	2656.4	72.5	3637	13	US-10-206-030-3	Sequence 3, Appl
C 17	2656.4	72.5	3637	15	US-10-211-079-3	Sequence 3, Appl
C 18	2656.4	72.5	3637	16	US-10-356-708-31	Sequence 31, Appl
C 19	2656.4	72.5	3637	16	US-10-205-772-3	Sequence 31, Appl
C 20	2656.4	72.5	3637	18	US-10-280-913A-31	Sequence 31, Appl
C 21	2656.4	72.5	3637	18	US-10-684-134-31	Sequence 31, Appl
C 22	2656.4	72.5	3637	18	US-10-637-758-31	Sequence 31, Appl
C 23	2609.4	71.3	4001	11	US-09-860-763-13	Sequence 13, Appl
C 24	2484.6	67.8	3774	11	US-09-860-763-9	Sequence 9, Appl
C 25	2469.4	67.4	5277	16	US-10-412-872-16	Sequence 16, Appl
C 26	2464.6	67.3	3908	17	US-10-014-099F-70	Sequence 70, Appl
C 27	2451.8	67.0	3928	15	US-10-128-590-45	Sequence 45, Appl
C 28	2451.8	67.0	3928	16	US-10-128-587A-45	Sequence 45, Appl
C 29	2451.8	67.0	3928	17	US-10-128-578B-45	Sequence 45, Appl
C 30	2446.8	66.8	3715	11	US-09-860-763-11	Sequence 11, Appl
C 31	2446.8	66.8	3927	17	US-10-014-099F-71	Sequence 71, Appl
C 32	2279.2	62.2	4119	9	US-09-993-170-1	Sequence 1, Appl
C 33	2271.2	62.0	4059	18	US-10-655-702-1	Sequence 1, Appl
C 34	2230.6	60.9	5250	18	US-10-474-070-51	Sequence 51, Appl
C 35	2216.2	60.5	4205	10	US-09-883-848A-48	Sequence 48, Appl
C 36	2216.2	60.5	4205	19	US-10-652-686-48	Sequence 48, Appl
C 37	2213	60.4	4205	10	US-09-883-848A-35	Sequence 35, Appl
C 38	2213	60.4	4205	19	US-10-652-686-35	Sequence 35, Appl
C 39	2172.8	59.3	7803	17	US-10-014-099F-82	Sequence 82, Appl
C 40	2172.8	59.3	8167	17	US-10-014-099F-83	Sequence 83, Appl
C 41	2172.6	59.3	10600	16	US-10-356-708-1	Sequence 1, Appl
C 42	2172.6	59.3	10600	18	US-10-280-913A-1	Sequence 1, Appl
C 43	2172.6	59.3	10600	18	US-10-684-134-1	Sequence 1, Appl
C 44	2172.6	59.3	10600	18	US-10-637-758-1	Sequence 1, Appl
C 45	2172.6	59.3	10624	16	US-10-356-708-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-10-001-189-41
; Sequence 41, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 3662
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac
; OTHER INFORMATION: sequence
; US-10-001-189-41

Query Match 100.0%; Score 3662; DB 13; Length 3662;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTAATGTAAGCGTTAAATATTTTGTAAATAATTCGCGTTAAATTTTGTAAATCAGCTC 60

Db 1 CTAATTTGTAAGCGTTAAATATTTTGTGTAATAATTCGCGTTAAATTTTGTGTAATAATCAGCTC 60
Qy 61 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAATCAAAAGAATAGACCGA 120
Db 61 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAATCAAAAGAATAGACCGA 120
Qy 121 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAAGTGGACTC 180
Db 121 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAAGTGGACTC 180
Qy 181 CAAAGCTCAAGAGGCGAAACCCGTCCTATCAGGGCGATGGCCCACTACGTGAACCAATCACC 240
Db 181 CAAAGCTCAAGAGGCGAAACCCGTCCTATCAGGGCGATGGCCCACTACGTGAACCAATCACC 240
Qy 241 CTAATCAAGTTTGTGGGTCGAGTGCAGTGCCTTAAGCACTAAATCGGAACCCCTAAAGGGAG 300
Db 241 CTAATCAAGTTTGTGGGTCGAGTGCCTTAAGCACTAAATCGGAACCCCTAAAGGGAG 300
Qy 301 CCCCAGATTTAGAGCTTGACGGGGAAGCCGGCAACGTGGCGAGAAAGGAAGGAAGAA 360
Db 301 CCCCAGATTTAGAGCTTGACGGGGAAGCCGGCAACGTGGCGAGAAAGGAAGGAAGAA 360
Qy 361 AGCGAAAGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGGTTAACCCAC 420
Db 361 AGCGAAAGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGGTTAACCCAC 420
Qy 421 CACACCGCGCGCTTAATGCGCCGTACAGGGCGGTCCCATTCGCCATTACAGCTGGG 480
Db 421 CACACCGCGCGCTTAATGCGCCGTACAGGGCGGTCCCATTCGCCATTACAGCTGGG 480
Qy 481 CAACTGTTGGGAAGGCGATCGGTGCGGCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
Db 481 CAACTGTTGGGAAGGCGATCGGTGCGGCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
Qy 541 GGGATGTGTCGAAGCGGATTAAGTTGGGTAAAGCGCAGGGTTTTCCCAAGTACAGCGTTG 600
Db 541 GGGATGTGTCGAAGCGGATTAAGTTGGGTAAAGCGCAGGGTTTTCCCAAGTACAGCGTTG 600
Qy 601 TAAACGAGCGGCGAGTGGCGGCTCGTTGATTCAGTTTGTGAACCGTGGAGACGG 660
Db 601 TAAACGAGCGGCGAGTGGCGGCTCGTTGATTCAGTTTGTGAACCGTGGAGACGG 660
Qy 661 GCAGACTCGCGTGCAATGTGTTTTACAGCGTGAAGAGATGAAGATCTCGACAC 720
Db 661 GCAGACTCGCGTGCAATGTGTTTTACAGCGTGAAGAGATGAAGATCTCGACAC 720
Qy 721 GCTGAGAACACGCGAGCTAGATTAAACCTTAGAAGATAATCATATTGTGACGTACGTTAA 780
Db 721 GCTGAGAACACGCGAGCTAGATTAAACCTTAGAAGATAATCATATTGTGACGTACGTTAA 780
Qy 781 AGATAATCATGCGTAAATTTGACGATGGATCTGTAATACGACTCACTATAGGGCGAAT 840
Db 781 AGATAATCATGCGTAAATTTGACGATGGATCTGTAATACGACTCACTATAGGGCGAAT 840
Qy 841 TGGGTACCGGGCCCCCTCGAGTTCGAGGTATCGATAAGCTTCGATTCGAAATTCCTGC 900
Db 841 TGGGTACCGGGCCCCCTCGAGTTCGAGGTATCGATAAGCTTCGATTCGAAATTCCTGC 900
Qy 901 AGCCCGGGGATCCACTAGTTCTAGAGCGGGCGCCACCGCGTGGAGCTCCAGCTTTTGT 960
Db 901 AGCCCGGGGATCCACTAGTTCTAGAGCGGGCGCCACCGCGTGGAGCTCCAGCTTTTGT 960
Qy 961 TCCCTTTAGTGGGTTAAATTAGATCCCATGCGTCAATTTTACGAGACTATCTTTCTAG 1020
Db 961 TCCCTTTAGTGGGTTAAATTAGATCCCATGCGTCAATTTTACGAGACTATCTTTCTAG 1020
Qy 1021 GGTAAATCTAGCTGATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCAATCGC 1080
Db 1021 GGTAAATCTAGCTGATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCAATCGC 1080
Qy 1081 CAAAGCTGCGCTATCTGGGCATCGGGGAGGAAGACCGCGTCCCTTTTCCCGCAGGTTG 1140

Db 1081 CAAAGCTGCGCTATCTGGGCATCGGGGAGGAAGACCGTGCCTTTTCCCGCAGGTTG 1140
Qy 1141 AAGCGGCATGGAAGAGATTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAC 1200
Db 1141 AAGCGGCATGGAAGAGATTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAC 1200
Qy 1201 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCCATGCAAGCCCTTAAACGGTGAACCTG 1260
Db 1201 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCCATGCAAGCCCTTAAACGGTGAACCTG 1260
Qy 1261 TCGTTTCAGGCCACCTGGGATACCACTTCGTCGGCTTTTCCGGACACAGTTCCGGATGG 1320
Db 1261 TCGTTTCAGGCCACCTGGGATACCACTTCGTCGGCTTTTCCGGACACAGTTCCGGATGG 1320
Qy 1321 TCAGCCCAAGCGCATCAGCAACCCGAAACATACCGCGACAGCCGGAACCTGCCGTGCCG 1380
Db 1321 TCAGCCCAAGCGCATCAGCAACCCGAAACATACCGCGACAGCCGGAACCTGCCGTGCCG 1380
Qy 1381 GTGTGCAGATTAAATGACAGCGGTGCGCGCTGGGATATTAAGTACAGCGAGGACGGGTATC 1440
Db 1381 GTGTGCAGATTAAATGACAGCGGTGCGCGCTGGGATATTAAGTACAGCGAGGACGGGTATC 1440
Qy 1441 CTGGCTGATGCGCAGAAATGGAATGGAATACCCCGTGAAGTTACCCGCGCGCGCTTT 1500
Db 1441 CTGGCTGATGCGCAGAAATGGAATGGAATACCCCGTGAAGTTACCCGCGCGCGCTTT 1500
Qy 1501 GCGTAAATCATGCTCATAGCTGTTTCCCTGTGTGAATTTGTTATCCGCTCACAAATTCACCA 1560
Db 1501 GCGTAAATCATGCTCATAGCTGTTTCCCTGTGTGAATTTGTTATCCGCTCACAAATTCACCA 1560
Qy 1561 CAAATACAGCGCGGAAAGCAATAAGTGTAAAGCCCTGGGGTGCCTAATGAGTGAGCTAACT 1620
Db 1561 CAAATACAGCGCGGAAAGCAATAAGTGTAAAGCCCTGGGGTGCCTAATGAGTGAGCTAACT 1620
Qy 1621 CACATTAATGCGTTGCGCTCACTGCGCGCTTTCAGTTCGGGAAACCTGTGTCGCCAGCT 1680
Db 1621 CACATTAATGCGTTGCGCTCACTGCGCGCTTTCAGTTCGGGAAACCTGTGTCGCCAGCT 1680
Qy 1681 GCATTAATGAATTCGGCAACCGCGGGGAGAGGGGTTTGGTATTGGGCGCTCTTCGCG 1740
Db 1681 GCATTAATGAATTCGGCAACCGCGGGGAGAGGGGTTTGGTATTGGGCGCTCTTCGCG 1740
Qy 1741 TTCTCTCGCTCACTGACTCGCTCGCTCGTTCGCTCGGCTCGGCGAGCGGTATCAGCTCA 1800
Db 1741 TTCTCTCGCTCACTGACTCGCTCGCTCGTTCGCTCGGCTCGGCGAGCGGTATCAGCTCA 1800
Qy 1801 CTCAAAGCGGTATAATACAGATTATCCAGAGATCAGGGATTAACGCAAGAAACATGTG 1860
Db 1801 CTCAAAGCGGTATAATACAGATTATCCAGAGATCAGGGATTAACGCAAGAAACATGTG 1860
Qy 1861 AGCAAAAGCGCAGCAAAAGGCGCAGGAACCGTAAAGGCGCGGTTGCTGGCGTTTTCCTCA 1920
Db 1861 AGCAAAAGCGCAGCAAAAGGCGCAGGAACCGTAAAGGCGCGGTTGCTGGCGTTTTCCTCA 1920
Qy 1921 TAGGCTCCGCCCTCGAGAGCATCAAAAAATCGACGCTCAAGTCAAGGTGCGGAAA 1980
Db 1921 TAGGCTCCGCCCTCGAGAGCATCAAAAAATCGACGCTCAAGTCAAGGTGCGGAAA 1980
Qy 1981 CCCGACAGGACTATAAAGATACAGGCGGTTTCCCGCTGGAAGCTCCCTGTCGCTCTCC 2040
Db 1981 CCCGACAGGACTATAAAGATACAGGCGGTTTCCCGCTGGAAGCTCCCTGTCGCTCTCC 2040
Qy 2041 TGTTCGAGCCCTGCGCTTACCGGATACCTGTCGCGCTTTTCCCTTCGGGAAGCGTGC 2100
Db 2041 TGTTCGAGCCCTGCGCTTACCGGATACCTGTCGCGCTTTTCCCTTCGGGAAGCGTGC 2100
Qy 2101 GCTTTCTCATAGCTCACTGATAGTATCTCAGTTTCGGTGTAGGTGCTTCGCTCCAAGCT 2160
Db 2101 GCTTTCTCATAGCTCACTGATAGTATCTCAGTTTCGGTGTAGGTGCTTCGCTCCAAGCT 2160
Qy 2161 GGGCTGTGTGACGAACCCCGGTTACGCCGACCGCTGCGGCTTATCGGTAACTATCG 2220
Db 2161 GGGCTGTGTGACGAACCCCGGTTACGCCGACCGCTGCGGCTTATCGGTAACTATCG 2220

```
QY 2221 TCTTGAGTCCAAACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAG 2280
Db 2221 TCTTGAGTCCAAACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAG 2280
QY 2281 GATTAGCAGCAGAGGTATGTAGGCGGTGCTACAGAGTTCTTCAGTGGTGGCCCTAACTA 2340
Db 2281 GATTAGCAGCAGAGGTATGTAGGCGGTGCTACAGAGTTCTTCAGTGGTGGCCCTAACTA 2340
QY 2341 CGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 2400
Db 2341 CGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 2400
QY 2401 AAAAGAGTTGGTAGTCTTGATTCGGGCAAAACAAACCAACCGCTGGTAGCGGTGGTTTTT 2460
Db 2401 AAAAGAGTTGGTAGTCTTGATTCGGGCAAAACAAACCAACCGCTGGTAGCGGTGGTTTTT 2460
QY 2461 TGTTCGACAGCAGATTCACGGCAGAGAAAGAGATCTCAAGAGATCTCTTCATCTT 2520
Db 2461 TGTTCGACAGCAGATTCACGGCAGAGAAAGAGATCTCAAGAGATCTCTTCATCTT 2520
QY 2521 TTCTACGGGCTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGATTTTGGTCATGAG 2580
Db 2521 TTCTACGGGCTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGATTTTGGTCATGAG 2580
QY 2581 ATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAAATGAAGTTTAAATCAAT 2640
Db 2581 ATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAAATGAAGTTTAAATCAAT 2640
QY 2641 CTAAAGTATATAGTAAGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTAGGCCACC 2700
Db 2641 CTAAAGTATATAGTAAGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTAGGCCACC 2700
QY 2701 TATCTCAGGATCTGTCTATTTTGGTTCATCTCATAGTTGCTGACTCCCGTGGTAGAT 2760
Db 2701 TATCTCAGGATCTGTCTATTTTGGTTCATCTCATAGTTGCTGACTCCCGTGGTAGAT 2760
QY 2761 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCCGAGACCC 2820
Db 2761 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCCGAGACCC 2820
QY 2821 AGCTCACCGGCTCCAGATTTATCAGCAATAACAGCCAGCCGAGGCGGCGAG 2880
Db 2821 AGCTCACCGGCTCCAGATTTATCAGCAATAACAGCCAGCCGAGGCGGCGAG 2880
QY 2881 AAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAGAGCTAG 2940
Db 2881 AAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAGAGCTAG 2940
QY 2941 AGTAAGTATGTCGCCAGTTPAATAGTTTGGCCAAACGTTGTGCAATTTGCTACAGGCATCGT 3000
Db 2941 AGTAAGTATGTCGCCAGTTPAATAGTTTGGCCAAACGTTGTGCAATTTGCTACAGGCATCGT 3000
QY 3001 GGTGTACGCTCGTGGTGGTATGGCTTCATTCAGTCCGTTCCCAACGATCAAGGG 3060
Db 3001 GGTGTACGCTCGTGGTGGTATGGCTTCATTCAGTCCGTTCCCAACGATCAAGGG 3060
QY 3061 AGTTACATGATCCCCATGTTGCAAAAAGCGTTAGCTCTCTCGGTCCTCCGATCGT 3120
Db 3061 AGTTACATGATCCCCATGTTGCAAAAAGCGTTAGCTCTCTCGGTCCTCCGATCGT 3120
QY 3121 TGTGAGAAGTAAGTTGGCCGAGTGTATCACTCATGTTTATGGCAGCAGCTGCATAATTC 3180
Db 3121 TGTGAGAAGTAAGTTGGCCGAGTGTATCACTCATGTTTATGGCAGCAGCTGCATAATTC 3180
QY 3181 TCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGATCGTGTAGTACTCAACCAAGTC 3240
Db 3181 TCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGATCGTGTAGTACTCAACCAAGTC 3240
QY 3241 ATTCTGAGAAATAGTATCGCGGACCGAGTTGCTCTTCGCCGGGTCAATACGGGATAA 3300
Db 3241 ATTCTGAGAAATAGTATCGCGGACCGAGTTGCTCTTCGCCGGGTCAATACGGGATAA 3300
```

```
QY 3301 TACCGGCGCACATAGCAGAACTTTAAAGTGTCTCATCTTGGAAAAAGTCTTTCGGGGCG 3360
Db 3301 TACCGGCGCACATAGCAGAACTTTAAAGTGTCTCATCTTGGAAAAAGTCTTTCGGGGCG 3360
QY 3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGAACCCACTCGTCACC 3420
Db 3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGAACCCACTCGTCACC 3420
QY 3421 CAACTGATCTTCAGCATCTTTTACTTCCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480
Db 3421 CAACTGATCTTCAGCATCTTTTACTTTCACAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480
QY 3481 GCAAAATGCGCAAAAAGGGAATAAGCGGCACACGGAATGTTGAATCTCATCTCTT 3540
Db 3481 GCAAAATGCGCAAAAAGGGAATAAGCGGCACACGGAATGTTGAATCTCATCTCTT 3540
QY 3541 CTTTTTCAATATATTGAAGCATTTATCAGGGTTATCTCATGCGGATACATATT 3600
Db 3541 CTTTTTCAATATATTGAAGCATTTATCAGGGTTATCTCATGCGGATACATATT 3600
QY 3601 TGAATGTATTTAGAAAAATAAACAATAGGGTTCGCGCACATTTCCCGAAAAAGTGCC 3660
Db 3601 TGAATGTATTTAGAAAAATAAACAATAGGGTTCGCGCACATTTCCCGAAAAAGTGCC 3660
QY 3661 AC 3662
Db 3661 AC 3662
```

RESULT 2

```
US-10-211-079-18/c
; Sequence 18, Application US/10211079
; Publication No. US20030148315A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Vaezhongs, Andrew A.
; APPLICANT: Vojdani, Fakhrish S.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding CEL I
; TITLE OF INVENTION: Endonuclease and Methods of Use Thereof
; FILE REFERENCE: P-LG 5381
; CURRENT APPLICATION NUMBER: US/10/211, 079
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 10/098,155
; PRIOR FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-211-079-18
```

```
Query Match 72.8%; Score 2665.4; DB 15; Length 3637;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;
```

```
QY 1 CTAATTTGTAACGGTTAATAATTTTGTAAAAATTCGGTTAAATTTTGTAAATCAGCTC 60
Db 3637 CTAATTTGTAACGGTTAATAATTTTGTAAAAATTCGGTTAAATTTTGTAAATCAGCTC 3578
QY 61 ATTTTAAACCAATAGGCGGAATCGCAAAATCCCTTAATAATCAAAAGATAGACCGA 120
Db 3577 ATTTTAAACCAATAGGCGGAATCGCAAAATCCCTTAATAATCAAAAGATAGACCGA 3518
QY 121 GATAGGGTTGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTGGACTC 180
Db 3517 GATAGGGTTGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTGGACTC 3458
QY 181 CAACTGTAAGGGCGGAAAAACCGTCTATCAGGGCGATGSCCCCACTACGTGAACCATCACC 240
Db 181 CAACTGTAAGGGCGGAAAAACCGTCTATCAGGGCGATGSCCCCACTACGTGAACCATCACC 240
```

Db	3457	CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCAC	3398
Qy	241	CTAATCAAGTTTTTTGGGGTCGAGTGGCGTAAAGCACTAAATCGGAACCCCTAAAGGGAG	300
Db	3397	CTAATCAAGTTTTTTGGGGTCGAGTGGCGTAAAGCACTAAATCGGAACCCCTAAAGGGAG	3338
Qy	301	CCCCGATTTAGAGCTTGACCGGGGAAGCCGCGCAACGTGGCGAGAAAAAGGAAGAA	360
Db	3337	CCCCCGATTTAGAGCTTGACCGGGGAAGCCGCGCAACGTGGCGAGAAAAAGGAAGAA	3278
Qy	361	AGCGAAAGAGCGGGCGCTAGGGCGCTGGCGAAGTGTAGGGTCAGCTGCGGGTAACAC	420
Db	3277	AGCGAAAGAGCGGGCGCTAGGGCGCTGGCGAAGTGTAGGGTCAGCTGCGGGTAACAC	3218
Qy	421	CACACCCCGCGCTTAATCGCGCGCTACAGGGCGCTCCCATTCGCCATTCAGGCTCG	480
Db	3217	CACACCCCGCGCTTAATCGCGCGCTACAGGGCGCTCCCATTCGCCATTCAGGCTCG	3158
Qy	481	CAACTGTTGGGAAGGGCGATCGGTGGGGCCCTCTTCGCTATTACGCCAGCTGGCGAAAG	540
Db	3157	CAACTGTTGGGAAGGGCGATCGGTGGGGCCCTCTTCGCTATTACGCCAGCTGGCGAAAG	3098
Qy	541	GGGATGCTGCGAAGCGGATTAAGTTGGGTAAAGCCAGGGTTTTCCAGTCACGACGTTG	600
Db	3097	GGGATGCTGCGAAGCGGATTAAGTTGGGTAAAGCCAGGGTTTTCCAGTCACGACGTTG	3038
Qy	601	TAAACAGCGCCAGTGAGCGCGCTCGTTCAATTCACCGTTTTTGAACCGGTGGAGGACGG	660
Db	3037	TAAACAGCGCCAGTGAGCGCGCTCGTTCAATTCACCGTTTTTGAACCGGTGGAGCTCA	2978
Qy	661	GCAGACTCGGGTGCAAAATGTTTTACAGGTGATGGAGCAGATGAAGATGCTCGAC	720
Db	2977	CGCGGTGGCGCGCTCTAGAACTAGTGGATCCCGCGGCTGCAGGAATCTTATTGTT	2918
Qy	721	GCTGCAGAACCGCAGCTAGATTAACCCCTAGAAAGATTAATCATATTGTGACGTAGTTAA	780
Db	2917	ATAGTTTCAATCCATGTCATGTGTAAATCCAGCAGCAGTTTACAACTCAAGAGGACCATGT	2858
Qy	781	AGATAATC-ATGCGTAAATTTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAA	839
Db	2857	GGTCACGCTTTTCGTTGGGATCTTTTCGAAAGGGCAGATTGTGTCACAGGTAATGGTTGT	2798
Qy	840	TTGGGTACCGGGCCCCCTCGAGGTGACGGTATCGATAAGCTTTGATATCGAATTCCTG	899
Db	2797	CTGGTAAAGGACAGGGCCATCGCCAAATGGAGTATTTTGTGAATGGTCTGCTAGTT	2738
Qy	900	CAGCCCGGGGATCACTAGTTCTAGAGGGCGCCACCGGGTGGAGCTCAGCTTTTG	959
Db	2737	GAACGGATCCATCTTTCAATGTTGTGGCGAAATTTTGAAGTTAGCTTTGATTCATTTCTTT	2678
Qy	960	TTCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTA	1019
Db	2677	GTTTGTCTGCCGTGATATACATTTGTGTAGTTTAAAGTTGTACTCGAGTTTGTGTCGA	2618
Qy	1020	GGGTTAACTAGCTGCATCAGGATCATATCGTGGGTCTTTTTTCCGGCTCAGTCATCGC	1079
Db	2617	GAAAGTTTCCATCTTCTTTAAATCAATACCTTTTAACTCGATAC- ----GATTAACAAG	2563
Qy	1080	CAAAGCTGGGCTATCTGGGCATTCGGGAGGAAGAACCGCGTCTTTTCCGGCGAGTT	1139
Db	2562	GGTATCACCTTCAAACTTGACTTTCAGCACGGGTCTTGTAGTTTCCCGTCACTTTTGAAGA	2503
Qy	1140	GAAGCGCATGGAAGATTTCGCGAGGATGACTGCTGCTGCAATTTGAGCTTTCAGCGAAGA	1199
Db	2502	TATAGTGGTTCTTGTAATTAACCTTCGGGATGCGCACTCTTTGAAAAAGTCAATGCCGTTT	2443
Qy	1200	CGCAGTTTACCATGATGATTTCGGGAAGGTGTGGCCATGCAGCCCTTTTAAAGGGTGAAC	1259
Db	2442	CATATGATCCGGATAACGAGAAAGCAATGAACACCATGAGAGAAAGTAGTGACAAAGT	2383
Qy	1260	TTCGTTACGGCCACTCGGGATACAGTTTCGTCGGCGCTTTTCCGGACACAGTTCCGGATG	1319
Db	2382	TGGCCATGGAACAGGTAGTTTTTCCAGTAGTGCATAATAAATTTAAGTGTAAAGCTTTCCGTA	2323
Qy	1320	GTCAAGCCGAAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCGGAACTGCCGTGCC	1379
Db	2322	TGTAGCATCATTCTCACCTCTCCACTGACAGAAAAATTTTGGCCATTAAATCACCATC	2263
Qy	1380	GGTGTGAGATTAATGACAGCGGTGCGGCTGGGATATTACGTGACGAGAGCGGGTAT	1439
Db	2262	TAAATCAACAAGAAAT-----TGGGACAACTCCAGTGAAGAATTC	2224
Qy	1440	CTGGCTGGATGCCGAGAAATGGACATGATACCCCGTGAGTTACCCGGCGGCGCGCT	1499
Db	2223	TTCTCTTTTACTCATCGGTATCCAGCTTTTGTTCCTTTTGTGAGGGTTAATTGCGCGCT	2164
Qy	1500	TGGCGTAAATCATGTGTCATAGCTGTTTCTGTGTGAAATTTTATCCGCTCAAAATCCAC	1559
Db	2163	TGGCGTAAATCATGTGTCATAGCTGTTTCTGTGTGAAATTTTATCCGCTCAAAATCCAC	2104
Qy	1560	ACAACATACGAGCGGAGCAATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAAGCTAAC	1619
Db	2103	ACAACATACGAGCGGAGCAATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAAGCTAAC	2044
Qy	1620	TCACATTAATTTGCGTTGCGCTCACTGCCCCGCTTTCCAGTCGGGAAACCTGTCGTCACAG	1679
Db	2043	TCACATTAATTTGCGTTGCGCTCACTGCCCCGCTTTCCAGTCGGGAAACCTGTCGTCACAG	1984
Qy	1680	TGCATTTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTCGCTATTGGGGGCTCTTCGG	1739
Db	1983	TGCATTTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTCGCTATTGGGGGCTCTTCGG	1924
Qy	1740	CTTCTCGCTCACTGACTGCTGCGTGGTTCGCTGGTGGTGGGCGAGCGGTATCAGCTC	1799
Db	1923	CTTCTCGCTCACTGACTGCTGCGTGGTTCGCTGGTGGTGGGCGAGCGGTATCAGCTC	1864
Qy	1800	ACTCAAGGCGGTAAATCGGTTATCCACAGAAATCAGGGGATAACGACAGAAAGACATGT	1859
Db	1863	ACTCAAGGCGGTAAATCGGTTATCCACAGAAATCAGGGGATAACGACAGAAAGACATGT	1804
Qy	1860	GAGCAAAAGGCGCAGCAAAAGCGCAGGAACCGCTAAAGGCGCGCTTGTGCGGTTTTTCC	1919
Db	1803	GAGCAAAAGGCGCAGCAAAAGCGCAGGAACCGCTAAAGGCGCGCTTGTGCGGTTTTTCC	1744
Qy	1920	ATAGGCTCCGCGCCCTGACAGCATCAAAATAATCGACGCTCAAGTCAGAGGTGGCGAA	1979
Db	1743	ATAGGCTCCGCGCCCTGACGAGCATCAAAATAATCGACGCTCAAGTCAGAGGTGGCGAA	1684
Qy	1980	ACCGACAGGACTAAAGATACAGAGGCTTTCCCTCGGAAGCTCCCTGTCGCTCTC	2039
Db	1683	ACCGACAGGACTAAAGATACAGAGGCTTTCCCTCGGAAGCTCCCTGTCGCTCTC	1624
Qy	2040	CTGTTCCGACCTCGCGCTTACCGGATACCTGTGCGCCTTTCTCCCTTCGGGAAGCGTGG	2099
Db	1623	CTGTTCCGACCTCGCGCTTACCGGATACCTGTGCGCCTTTCTCCCTTCGGGAAGCGTGG	1564
Qy	2100	CGTTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTTCGTTGTTAGTTCGTTCCCAAGC	2159
Db	1563	CGTTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTTCGTTGTTAGTTCGTTCCCAAGC	1504
Qy	2160	TGGGCTGTGCGACGAACCCCGTTTACGCGCGAGCGCTGCGCTTATCCGCTAACTATC	2219
Db	1503	TGGGCTGTGCGACGAACCCCGTTTACGCGCGAGCGCTGCGCTTATCCGCTAACTATC	1444
Qy	2220	GTCTTGAGTCCAAACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA	2279
Db	1443	GTCTTGAGTCCAAACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA	1384
Qy	2280	GGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCCTAACT	2339
Db	1383	GGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCCTAACT	1324
Qy	2340	ACGGCTACCTAGAGGACAGTATTTGGTATCTCGGCTCTGCTGAAGCAGTTACTCTCG	2399
Db	1323	ACGGCTACCTAGAGGACAGTATTTGGTATCTCGGCTCTGCTGAAGCAGTTACTCTCG	1264

QY 2400 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCAACCGCTGGTAGCGGTGTTTTT 2459
DB 1263 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCGCTGGTAGCGGTGTTTTT 1204
QY 2460 TTGTTTGCAGCAGCAGATAGCGCAGAAAAAGGATCTCAAGAAAGATCTTTTGATCT 2519
DB 1203 TTGTTTGCAGCAGCAGATAGCGCAGAAAAAGGATCTCAAGAAAGATCTTTTGATCT 1144
QY 2520 TTCTACCGGGTCTGACGCTCAGTGGAACGAAACCTCACGTTAAAGGATTTTGGTCATGA 2579
DB 1143 TTCTACCGGGTCTGACGCTCAGTGGAACGAAACCTCACGTTAAAGGATTTTGGTCATGA 1084
QY 2580 GATTATCAAAAAGGATCTTACCTAGATCTTTTAAATTTAAATTTAAATTTAAATCAA 2639
DB 1083 GATTATCAAAAAGGATCTTACCTAGATCTTTTAAATTTAAATTTAAATTTAAATCAA 1024
QY 2640 TCTAAAGTATATAGTAAATCTTGGTCTGACAGTTTACCAATGCTTAAATCAGTAGGCAC 2699
DB 1023 TCTAAAGTATATAGTAAATCTTGGTCTGACAGTTTACCAATGCTTAAATCAGTAGGCAC 964
QY 2700 CTATCTCAGCGATCTGTCTATTTCGTTTCATCATAGTTGCTGACTCCCGTGTGTAGA 2759
DB 963 CTATCTCAGCGATCTGTCTATTTCGTTTCATCATAGTTGCTGACTCCCGTGTGTAGA 904
QY 2760 TAACTACGATCGGAGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGCGAGAC 2819
DB 903 TAACTACGATCGGAGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGCGAGAC 844
QY 2820 CACGCTCACCGCTCCAGATTTATCAGCAATTAACCCAGCAGCGGAGGCGGAGCGCA 2879
DB 843 CACGCTCACCGCTCCAGATTTATCAGCAATTAACCCAGCAGCGGAGGCGGAGCGCA 784
QY 2880 GAAGTGGTCTGCAATTTATCCGCTCCATCCAGTCTTAAATTTGTTCCGGGAAGCTA 2939
DB 783 GAAGTGGTCTGCAATTTATCCGCTCCATCCAGTCTTAAATTTGTTCCGGGAAGCTA 724
QY 2940 GAGTAAGTAGTTCCGAGTTAATAGTTTCCGCAAGCTTGTGCTTGCATTCAGGCATCG 2999
DB 723 GAGTAAGTAGTTCCGAGTTAATAGTTTCCGCAAGCTTGTGCTTGCATTCAGGCATCG 664
QY 3000 TCGTGTACGCTCGTCTGTTGGTATGGCTTCATTACAGCTCCGTTCCCAACGATCAAGC 3059
DB 663 TCGTGTACGCTCGTCTGTTGGTATGGCTTCATTACAGCTCCGTTCCCAACGATCAAGC 604
QY 3060 GAGTTACATGATCCCCCATGTTGTGCAAAAAGCGTTAGCTCTTCCGTCCTCCGATCG 3119
DB 603 GAGTTACATGATCCCCCATGTTGTGCAAAAAGCGTTAGCTCTTCCGTCCTCCGATCG 544
QY 3120 TTGTCAAGAGTAAGTTGGCCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATT 3179
DB 543 TTGTCAAGAGTAAGTTGGCCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATT 484
QY 3180 CTCCTTACTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGGTAGTACTCAACCAAGT 3239
DB 483 CTCCTTACTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGGTAGTACTCAACCAAGT 424
QY 3240 CATTTCTGAGATAGTGTATGCGCCGACCGAGTTGCTCTTGGCCGGCGTCAATACGGGATA 3299
DB 423 CATTTCTGAGATAGTGTATGCGCCGACCGAGTTGCTCTTGGCCGGCGTCAATACGGGATA 364
QY 3300 ATACCGGCCACATAGCAGAACTTTTAAAGTGTCTATCATTTGGAACAGTTCTTCGGGCG 3359
DB 363 ATACCGGCCACATAGCAGAACTTTTAAAGTGTCTATCATTTGGAACAGTTCTTCGGGCG 304
QY 3360 GAAACTCTCAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCACTCGTGCAC 3419
DB 303 GAAACTCTCAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCACTCGTGCAC 244
QY 3420 CCAACTGATCTTTCAGCATCTTTTACCTTTTCCAGCGTTTCTGGGTGAGCAAAAACAGGAA 3479
DB 243 CCAACTGATCTTTCAGCATCTTTTACCTTTTCCAGCGTTTCTGGGTGAGCAAAAACAGGAA 184
QY 3480 GGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAAATACTCATCTCT 3539

DB 183 GGCAAAATCCCGCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAAATCTCATCTCT 124
QY 3540 TCCTTTTCAATATTATTGAAGCAATTTATCAGGGTTATTGTCTCTCATGAGCGGATACATAT 3599
DB 123 TCCTTTTCAATATTATTGAAGCAATTTATCAGGGTTATTGTCTCTCATGAGCGGATACATAT 64
QY 3600 TTGAATGTATTAGAAAAATAAAACAAATAGGGGTTCCGGCGCACATTTCCCGAAAAAGTGC 3659
DB 63 TTGAATGTATTATTAGAAAAATAAAACAAATAGGGGTTCCGGCGCACATTTCCCGAAAAAGTGC 4
QY 3660 CAC 3662
DB 3 CAC 1

RESULT 3
US-10-356-708-32/c
; Publication 32, Application US/10356708
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: MISMATCH ENDONUCLEASES AND METHODS OF USE THEREOF
; FILE REFERENCE: P-LG 10100
; CURRENT APPLICATION NUMBER: US/10/356,708
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: 60/353,722
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 10/098,155
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 10/211,079
; PRIOR FILING DATE: 2002-08-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Encodes Cycle 3 BFP gene
US-10-356-708-32

Query Match 72.8%; Score 2665.4; DB 16; Length 3637;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;

QY 1 CTAATTTGTAACGCTTAATTTTGTAAAAATTCGGTTAAATTTTGTGTTAAATCAGTCT 60
DB 3637 CTAATTTGTAACGCTTAATTTTGTAAAAATTCGGTTAAATTTTGTGTTAAATCAGTCT 3578
QY 61 ATTTTAAACCAATAGCGCAAAATCGCAAAATCCCTTATAAATCAAAAAGATAGACCGA 120
DB 3577 ATTTTAAACCAATAGCGCAAAATCGCAAAATCCCTTATAAATCAAAAAGATAGACCGA 3518
QY 121 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTTATTAAGAACGTGACTC 180
DB 3517 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTTATTAAGAACGTGACTC 3458
QY 181 CAACGTCAAAAGGCGCAAAAACCGTCTATCAGGGCGATGCCCCACTACGTGAACCATCACC 240
DB 3457 CAACGTCAAAAGGCGCAAAAACCGTCTATCAGGGCGATGCCCCACTACGTGAACCATCACC 3398
QY 241 CTAATCAAGTTTTTGGGTCGAGGTGCGTAAAGCACTAAATCGGAACCCCTAAAGGAG 300
DB 3397 CTAATCAAGTTTTTGGGTCGAGGTGCGTAAAGCACTAAATCGGAACCCCTAAAGGAG 3338
QY 301 CCCCAGATTGAGCTTGAACGGGAAAGCCGGGAAACGTGGCGAGAAAGGAAGAA 360
DB 3337 CCCCAGATTGAGCTTGAACGGGAAAGCCGGGAAACGTGGCGAGAAAGGAAGAA 3278
QY 361 ACGGAAGAGCGGCGCTAGGCGCTGCAAGTGTAGCGGTACGCTCGCGTAAACAC 420
DB 3277 ACGGAAGAGCGGCGCTAGGCGCTGCAAGTGTAGCGGTACGCTCGCGTAAACAC 3218

QY 421 CACACCCGCGCTTAATGCGCGCTACAGGGCGCGTCCCATTCGCCATTGAGGCTCGG 480
Db |||||
3217 CACACCCGCGCGCTTAATGCGCGCTACAGGGCGCGTCCCATTCGCCATTGAGGCTCGG 3158
QY 481 CAACCTGTTGGGAAGGCGCATCGGTGCGGCGCTCTTTGCTATTAACGCCAGCTGGCGAAAGG 540
Db |||||
3157 CAACCTGTTGGGAAGGCGCATCGGTGCGGCGCTCTTTGCTATTAACGCCAGCTGGCGAAAGG 3098
QY 541 GGGATGCTCTCAAGGCGATTAAAGTTGGTAAAGCCAGGCTTTCCCAAGTCAAGGCTTG 600
Db |||||
3097 GGGATGCTCTCAAGGCGATTAAAGTTGGTAAAGCCAGGCTTTCCCAAGTCAAGGCTTG 3038
QY 601 TAAACAGCGCGCAGTGAAGCGCGCTCGTTCAATTCACCGTCTTTTGAACCCGTCGAGGACGG 660
Db |||||
3037 TAAACAGCGCGCAGTGAAGCGCGCTTAATACGACTCACTATAGGCGGAATTGGAGCTCA 2978
QY 661 GCAGACTCGGGTCAAAATGTTTAAAGCGGTGATGAGCAGATGAAGATGCTCGACAC 720
Db |||||
2977 CCGCGGTGGCGCGCTCTAGAACTAGTGGATCCCGCGGCTGCAGGAATCTTATTTGT 2918
QY 721 GCTGCAGAACGCGCAGCTAGATTAACTTAACTAGAAAGATATCATATTGAGCTAGCTTAA 780
Db |||||
2917 ATAGTTTCAATGCCATGTTGTAATCCCGAGCAGGTTTACAACTCAAGAAAGGACCATGT 2858
QY 781 AGATAATC-ATGCGGTAAATTAAGCATGCGATCTGTAAATACGACTCACTATAGGCGGAA 839
Db |||||
2857 GGTACGCTTTTCGTTGGGATCTTTCGAAAGGCGAGATTGTCGACAGGTAATGGTTGT 2798
QY 840 TTGGGTACCGGGCCCCCTCGAGGTGCAAGGTATCGATAGCTTGAATTCGAATTCCTG 899
Db |||||
2797 CTGGTAAAGGACAGGGCCATCGGCAATTTGGAGTATTTGTTGATAATGGTCTGCTAGTT 2738
QY 900 CAGCCCGGGGATCCACTAGTTCTAGAGCGCGCCAGCGGTGGAGCTCCAGCTTTTG 959
Db |||||
2737 GAACGGATCCACTTTCAAGTTTGGCGGAAATTTTGAAGTGTAGTTCATTCATTTT 2678
QY 960 TTCCTTTTGTAGGGTTTAAATAGATCCCATGCGTCAATTTTACGACAGCTATCTTTCTA 1019
Db |||||
2677 GTTTGCTGCGGTGATGATATACATTTGTGAGTTTAAAGTTGTACTCGAGTTTGTCCGA 2618
QY 1020 GGGTTAATCTAGCTGCATCAGATCATATCGTCGGGTCTTTTTCCGGCTCAGTCATCGC 1079
Db |||||
2617 GAAATGTTTCCATCTTTTAAATCAATACCTTTTAACTCGATAC- ----GATTAAACAAG 2563
QY 1080 CCAAGCTGGCGTATCTCGGCATCGGGAGGAAGACCGGCTTTTCCCGGAGGT 1139
Db |||||
2562 GGTATCACCTTCAAACTTGACTTCAGACGGCTTTGTAGTTCGGCTCATCTTTGAAAGA 2503
QY 1140 GAAGCGCATGGAAAGAGTTTGGCGAGGATGACTGCTGCTGCATTTGACGTTGAGCGAAAA 1199
Db |||||
2502 TATAGTGGTTCTGTACATAAACCCTTCGGGATGSCACTTTGAAAAAAGTCATGCGGTTT 2443
QY 1200 CGCAGTTTACATGATGATTCGGGAAGGTGCGCATGACGCTTTTAAAGGTGAACG 1259
Db |||||
2442 CATATGATCCGGATAACGAGAAAAAGCATTTGAACACCATGAGAGAAAGTAGTGACAAAGTGT 2383
QY 1260 TTCGTTCAAGGCCACTGGGATACAGTTTCGTGCGGCTTTTCCGACACAGTTCGGATG 1319
Db |||||
2382 TGGCCATGGAACAGGTAGTTTTCAGTGTAGTGAATAAATTAAGTGTAAAGCTTTCCGTA 2323
QY 1320 GTCAGCCGGAAGCGCATCAGCAACCCGGAACAATACCGCGCAGACGCGGAACCTGCGTGCC 1379
Db |||||
2322 TGTAGCATCACTTCACCTCTCCATCTGACAGAAAAATTTGTGCGCATTTAAATCACCATC 2263
QY 1380 GGTGTGAGATTAATGACAGCGGTGCGGCTGGGATATTAGTCAGCGAGACGGGTAT 1439
Db |||||
2262 TAAATCAACAAGAAAT- -----TGGGCAAACTCCAGTGAAGATTC 2224
QY 1440 CTTGGTGGATGCCGCAAAATGGACATGATACCCCGTGTAGTTTACCCGGCGGCGCT 1499
Db |||||
2223 TTTCTCTTACTCATCGGTACCGAGCTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCT 2164

QY 1500 TGGCGTAATCATGTGCATAGCTGTTTCTGTGTGAAATTTTATCCGCTCAAAATTCAC 1559
Db |||||
2163 TGGCGTAATCATGTGCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCAAAATTCAC 2104
QY 1560 ACAACATACGAGCGGGAAGCATAAAGTGTAAAGCTTGGGTGCCCTAAATGAGTGAGCTAAC 1619
Db |||||
2103 ACAACATACGAGCGGGAAGCATAAAGTGTAAAGCTTGGGTGCCCTAAATGAGTGAGCTAAC 2044
QY 1620 TCACATTAATTTGCGTTGGCTCACTGCGCGCTTTTCCAGTCGGGAAACCTGTCGTCGACG 1679
Db |||||
2043 TCACATTAATTTGCGTTGGCTCACTGCGCGCTTTTCCAGTCGGGAAACCTGTCGTCGACG 1984
QY 1680 TGCATTAATGAATCGGCAACGCGCGGGAGAGCGGTTTCCGTAATTCGGCGCTTCTCG 1739
Db |||||
1983 TGCATTAATGAATCGGCAACGCGCGGGAGAGCGGTTTCCGTAATTCGGCGCTTCTCG 1924
QY 1740 CTTCTCGCTCACTGACTCGCTCGGCTCGGTCGTTCCGCTCGGGGAGCGGTATCAGCTC 1799
Db |||||
1923 CTTCTCGCTCACTGACTCGCTCGGCTCGGTCGTTCCGCTCGGGGAGCGGTATCAGCTC 1864
QY 1800 ACTCAAGGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGACGAAAGAACATGT 1859
Db |||||
1863 ACTCAAGGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGACGAAAGAACATGT 1804
QY 1860 GAGCAAAAGCCAGCAAAAGCCAGGAAACCGTAAAGGCGCGCTTGTGCGGTTTTTCC 1919
Db |||||
1803 GAGCAAAAGGCGCAGCAAAAGCCAGGAAACCGTAAAGGCGCGCTTGTGCGGTTTTTCC 1744
QY 1920 ATAGGCTCCGCGCGCTGACGAGCATCAAAAAATCAGAGCTCAAGTCAGAGGTGGCGAA 1979
Db |||||
1743 ATAGGCTCCGCGCGCTGACGAGCATCAAAAAATCAGAGCTCAAGTCAGAGGTGGCGAA 1684
QY 1980 ACCGACAGGACTATAAAGATACAGGGGTTTCCCGCTGGAAGCTCCCTCGTGCCTCTC 2039
Db |||||
1683 ACCGACAGGACTATAAAGATACAGGGGTTTCCCGCTGGAAGCTCCCTCGTGCCTCTC 1624
QY 2040 CTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAAGCGTG 2099
Db |||||
1623 CTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAAGCGTG 1564
QY 2100 GCGTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTCGCTCCCAAGC 2159
Db |||||
1563 GCGTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTCGCTCCCAAGC 1504
QY 2160 TGGCTGTGTGCAGCAACCCCGCTTCAGCCCGACCGCTGCGGCTTTATCCGCTAACTATC 2219
Db |||||
1503 TGGCTGTGTGCAGCAACCCCGCTTCAGCCCGACCGCTGCGGCTTTATCCGCTAACTATC 1444
QY 2220 GTCCTGAGTCCAAACCCGCTAAGACACGACTTATCGCACTGGCAGCAGCAGCTGGTAACA 2279
Db |||||
1443 GTCCTGAGTCCAAACCCGCTAAGACACGACTTATCGCCACTGGCAGCAGCAGCTGGTAACA 1384
QY 2280 GGATTCAGAGCGAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGTGGCTTAACCTCG 2339
Db |||||
1383 GGATTCAGAGCGAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGTGGCTTAACCTCG 1324
QY 2340 ACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAAGCGAGTTACTTCG 2399
Db |||||
1323 ACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCGAGTTACTTCG 1264
QY 2400 GAAAAAGAGTTGTAGTCTTTGATCCCGGCAAAACAAACCCGCTGTGTAGCGGTGTTTTT 2459
Db |||||
1263 GAAAAAGAGTTGTAGTCTTTGATCCCGCAAAACAAACCCGCTGTGTAGCGGTGTTTTT 1204
QY 2460 TTGTTTCCAGCAGCAGATTAACGCGCAGAAAAAAGGATCTCAAGAGATCCTTTGATCT 2519
Db |||||
1203 TTGTTTGAAGCGAGATTAACGCGCAGAAAAAAGGATCTCAAGAGATCCTTTGATCT 1144
QY 2520 TTTCTTACGGGCTGTGAGCTCAGTGGAAACGAAAACTCAGTTTAAAGGGATTTTGTGTCATGA 2579
Db |||||
1143 TTTCTTACGGGCTGTGAGCTCAGTGGAAACGAAAACTCAGCTTAAAGGGATTTTGTGTCATGA 1084
QY 2580 GATTATCAAAAAAGGATCTTACCTAGATCCTTTTAAATTAATAATGAAGTTTAAATCA 2639

Db 1083 GATTATCAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAAGTTTAAATCAA 1024
Qy TCTAAAGTATATAGTAAACTTGGTCTGACAGTTTACCAATGCTTAATCAGTGAGGCAC 2699
Db 1023 TCTAAAGTATATAGTAAACTTGGTCTGACAGTTTACCAATGCTTAATCAGTGAGGCAC 964
Qy CTATCTCAGCGATCTGTCTATTTTGGTTCATCCATAGTTGCTGCTGACCTCCCGTGTGTAGA 2759
Db 963 CTATCTCAGCGATCTGTCTATTTTGGTTCATCCATAGTTGCTGCTGACCTCCCGTGTGTAGA 904
Qy TAACTACGATACGGGAGGCTTACCATCTGGCCCGACGCTGCTCAATGATACCGCGAGACC 2819
Db 903 TAACTACGATACGGGAGGCTTACCATCTGGCCCGACGCTGCTCAATGATACCGCGAGACC 844
Qy CACGCTCACCGGCTCCAGATTTTATCAGCAATAAACCAGCAGCGCGGAAGCGCGACGCA 2879
Db 843 CACGCTCACCGGCTCCAGATTTTATCAGCAATAAACCAGCAGCGCGGAAGCGCGACGCA 784
Qy GAAGTGGTCTCGAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTCCGGGAAGCTA 2939
Db 783 GAAGTGGTCTCGAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTCCGGGAAGCTA 724
Qy GAGTAAGTAGTTCCGAGTTAATAGTTTGGCGCAAGTTGCTGCCATTTGCTACAGGCATCG 2999
Db 723 GAGTAAGTAGTTCCGAGTTAATAGTTTGGCGCAAGTTGCTGCCATTTGCTACAGGCATCG 664
Qy TGGTGTACGCTCGTCTGGTGTATGGCTTTCATTACAGCTCCGTTCCCAACGATCAAGCG 3059
Db 663 TGGTGTACGCTCGTCTGGTGTATGGCTTTCATTACAGCTCCGTTCCCAACGATCAAGCG 604
Qy GAGTTACATGATCCCCATTTGTGTGCAAAAAAGCGGTAGCTCTCTTCCGATCG 3119
Db 603 GAGTTACATGATCCCCATTTGTGTGCAAAAAAGCGGTAGCTCTCTTCCGATCG 544
Qy TTGTGAGAAGTAAGTTGGCGGAGTGTATCACTCATGCTTATGGCAGCACTGCATATTT 3179
Db 543 TTGTGAGAAGTAAGTTGGCGGAGTGTATCACTCATGCTTATGGCAGCACTGCATATTT 484
Qy CTCTTACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTACTCAACCACT 3239
Db 483 CTCTTACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTACTCAACCACT 424
Qy CATTCGAGAAATGATGTATGCGCGACGAGTTGCTCTTGGCCCGGCTCAATACGGGATA 3299
Db 423 CATTCGAGAAATGATGTATGCGCGACGAGTTGCTCTTGGCCCGGCTCAATACGGGATA 364
Qy ATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAACGTTTCTTGGGGC 3359
Db 363 ATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAACGTTTCTTGGGGC 304
Qy GAAAACTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCAACCACTCTGTGCAC 3419
Db 303 GAAAACTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCAACCACTCTGTGCAC 244
Qy CCAACTGATCTTTCAGCATCTTTTACCTTCCACGAGTCTTCTGGTGAGCAAAAACAGGAA 3479
Db 243 CCAACTGATCTTTCAGCATCTTTTACCTTCCACGAGTCTTCTGGTGAGCAAAAACAGGAA 184
Qy GGCAAAATGCCCAAAAAGGAATTAAGGCGACACGGAATTTGTAATCTCATCTCT 3539
Db 183 GGCAAAATGCCCAAAAAGGAATTAAGGCGACACGGAATTTGTAATCTCATCTCT 124
Qy TCCCTTTTCAATATTTTCAAGCATTTATCAGGTTTATGCTCATGAGCGGATACATAT 3599
Db 123 TCCCTTTTCAATATTTTCAAGCATTTATCAGGTTTATGCTCATGAGCGGATACATAT 64
Qy TTGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGCGACATTTTCCCGGAAAGTGC 3659
Db 63 TTGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGCGACATTTTCCCGGAAAGTGC 4
Qy 3660 CAC 3662
|||

Db 3 CAC 1
RESULT 4
US-10-280-913A-32/c
; Sequence 32, Application US/10280913A
; Publication No. US20040110130A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PIG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/280,913A
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Encodes Cycle 3 BFP gene
US-10-280-913A-32
Query Match 72.8%; Score 2665.4; DB 18; Length 3637;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;
Qy 1 CTAATTCGTAAAGCGTTAATATTTTGTAAAAATTCGGTTAAATTTTGTAAATCAGCTC 60
Db 3637 CTAATTCGTAAAGCGTTAATATTTTGTAAAAATTCGGTTAAATTTTGTAAATCAGCTC 3578
Qy 61 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 120
Db 3577 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 3518
Qy 121 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGACTC 180
Db 3517 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGACTC 3458
Qy 181 CAACTCAAGGCGGAAACCGTCTATCAGGGCGATGGCCCACTACGTTGAACCATCACC 240
Db 3457 CAACTCAAGGCGGAAACCGTCTATCAGGGCGATGGCCCACTACGTTGAACCATCACC 3398
Qy 241 CTAATCAAGTTTGTGGGTTCGAGGTGCGTAAAGCACTAAATCGGAAACCTTAAAGGGAG 300
Db 3397 CTAATCAAGTTTGTGGGTTCGAGGTGCGTAAAGCACTAAATCGGAAACCTTAAAGGGAG 3338
Qy 301 CCCCCGATTTAGAGCTTGACGGGGAAGCCGCGAACCGTGGCGAGAAAGAAAGGGAAGAA 360
Db 3337 CCCCCGATTTAGAGCTTGACGGGGAAGCCGCGAACCGTGGCGAGAAAGGGAAGAA 3278
Qy 361 AGCGAAAGAGCGGGCGCTAGGGCGCTGCAAGTGTAGCGGTACGCTGGCGGTACCCAC 420
Db 3277 AGCGAAAGAGCGGGCGCTAGGGCGCTGCAAGTGTAGCGGTACGCTGGCGGTACCCAC 3218
Qy 421 CACACCGCGCGCTTAAATGCGCGCTACAGGGCGGTCCCAATTCGCCATTTCAGGCTGCG 480
Db 3217 CACACCGCGCGCTTAAATGCGCGCTACAGGGCGGTCCCAATTCGCCATTTCAGGCTGCG 3158
Qy 481 CAACTGTTGGGAAGGCGGATCGGTGCGGCCCTCTTTCGCTATTACGCCAGCTGCGGAAAG 540
Db 3157 CAACTGTTGGGAAGGCGGATCGGTGCGGCCCTCTTTCGCTATTACGCCAGCTGCGGAAAG 3098
Qy 541 GGGATGCTGCAAGCGGATTAAGTTGGGTAAAGCGGAGGTTTCCAGTCAGGCTG 600
|||

Db 3097 GGGATGTGCTCAAGCGCGAATTAAGTTGGGTAAAGCCACGGGTTTTCCCACTCACGACGTTG 3038
Qy 601 TAAACACGACGCGCAGTGTAGCGCGCTCGTTCAATTCACGTTTTTGAACCCGCTGGAGACGG 660
Db 3037 TAAACACGACGCGCAGTGTAGCGCGCTGAATACGACTCATATAGGCGGAATTTGGAGCTCCA 2978
Qy 661 GCAGACTCGCGGTGCAAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
Db 2977 CCGCGGTGCGCGCGCTAGNACTAGTGGATCCCGCGGCTGCAGGAATTTCTTATTGTT 2918
Qy 721 GCTGCAGAACACGACGTAGATTAACCTTAGAAGATTAATCATATGTGTGACGTAGCTTAA 780
Db 2917 ATAGTTTCAATCCATGTCATGTGTAATCCCGACGACAGTTTACAACTCAAGAAAGGACCATGT 2858
Qy 781 AGATAATC-ATGCGTAAATTTGACGCAATGGGATCTGTAAATACGACTCACTATAGGGCGAA 839
Db 2857 GGTCAACGCTTTTCGTTGGGATCTTTTCGAAAGGGCAGATTTGTGTGACAGGTAATGGTTGT 2798
Qy 840 TTGGGTACCGGCGCCCGCTCGAGGTGCAGCGTATCGATAAGCTTTGATATCGAAATTCCTG 899
Db 2797 CTGGTAAAGGACAGGGCCATCGCCAAATTTGGAGTATTTTGTGATAATGGTCTGCTAGTT 2738
Qy 900 CAGCCCGGGGATCCACTAGTTTCTAGAGCGCGCCACCGCGGTGAGAGCTCCAGCTTTTGG 959
Db 2737 GAAAGGATCCATCTTTCAATGTGTGGCGAAATTTTGAAGTTAGCTTTGATTCATTTCTTTT 2678
Qy 960 TTCCCTTTAGTGAGGGTTAAATTTAGATCCATCGGTCAATTTTACGCACTATCTTTCTA 1019
Db 2677 GTTTGTCTGCGGTGATGATACATTTGTGTGAGTTTAAAGTTGTACTCGAGTTTGTGTCGA 2618
Qy 1020 GGGTTAACTTAGCTGATCAGGATCATATCGTCGGGTCTTTTTCCGGCTCAGTCATCGC 1079
Db 2617 GAAATGTTTCCATCTTCTTTTAAATCAATACCTTTTAACTTCGNATC-----GATTAAACAAG 2563
Qy 1080 CCAAGCTGGCGTATCTCGGCATCGGGGAGGAGAAAGCCCGTGCCTTTTCCCGGAGGTT 1139
Db 2562 GGTATCACCTTCAAACTTTGACTTTCAGCACGGGTCTTGTAGTTCCCGTCACTTTTGAAGA 2503
Qy 1140 GAAGCGCATGMAAGAGTTTGGCGAGGATGACTGCTGCTGCATTTGACGTTGAGCGAATA 1199
Db 2502 TATAGTGGGTCTCTGTAATCAATACCTTTCGGGATGGCACTCTTGAATAAAGTCATGCCGTTT 2443
Qy 1200 CGCACGTTTACCATGATGATTTCCGGAAGGTGTGGCCATGCACGCGCTTTTAAAGCGTGAAC 1259
Db 2442 CATATGATCCGGATAACAGAAAGCAATGAACACCATGAGAAAGTAGTGACAAGTGT 2383
Qy 1260 TTCGTTAGGCGCACCTGGGATACAGTTTCGTGCGGGCTTTTCCGGACACAGTTCCGGATG 1319
Db 2382 TGGCCATGGAACAGGTAGTTTTCAGTAGTGTGCAATAAAATTTAAGTGAAGCTTTCCGTA 2323
Qy 1320 GTCAGCCGAGCGCATCAGCAACCCGAAACATACCGGCGACAGCGGAACCTGCGTGCC 1379
Db 2322 TGTAGCATCACCTTACCCTCTCCACTGACAGAAATTTGTGCCCATTAACATCACCATC 2263
Qy 1380 GGTGTGAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTCAAGCGAGAGCGGTAT 1439
Db 2262 TAAATCAACAGAAAT-----TGGGACAACTCCAGTGAAAGTTT 2224
Qy 1440 CCTGGCTGGATGCGCGCAGAAATGGAACATGGAATACCCCGTAGTTTACCCGCGGGCGCGCT 1499
Db 2223 TTCTCTTTTACTCATCGGTACCCAGCTTTTGTTCCTTTTAGTGGGGTTAAATTTGCGCGCT 2164
Qy 1500 TGGCGTAACTAGTCATAGCTGTTTCTCTGTGGAATTTGATTCGCTCACAATTTCCAC 1559
Db 2163 TGGCGTAACTAATGGTTCATAGCTGTTTCTGTGTGAAATTTGTTATTCGCTCACAATTTCCAC 2104
Qy 1560 ACAACATACGAGCCGGAAGCATAAAGTGTAAAGCTCGGGTGCCTTAATGAGTGAGCTTAAC 1619
Db 2103 ACAACATACGAGCCGGAAGCATAAAGTGTAAAGCTCGGGTGCCTTAATGAGTGAGCTTAAC 2044
Qy 1620 TCACATTAATTTGCGTTGCGCTCATCTGCCGCTTTCCAGTCGGGAAACCTGTCGTCGACAG 1679
Db 2043 TCACATTAATTTGCGTTGCGCTCACTGCCCGCTTTTCCAGTCGGGAAACCTGTCGTCGACAG 1984

Qy 1680 TGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTTGCGTATTGGGGCTCTTCGG 1739
Db 1983 TGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTTGCGTATTGGGGCTCTTCGG 1924
Qy 1740 CTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1799
Db 1923 CTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1864
Qy 1800 ACTCAAGGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACCGCAGGAAGACATGT 1859
Db 1863 ACTCAAGGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACCGCAGGAAGACATGT 1804
Qy 1860 GAGCAAAAGGCGCAGCAAAAGCCAGGAACCGCTAAAAAGCCGCTGTGTGCGGTTTTTCC 1919
Db 1803 GAGCAAAAGGCGCAGCAAAAGCCAGGAACCGCTAAAAAGCCGCTGTGTGCGGTTTTTCC 1744
Qy 1920 ATAGGCTCCGCGCCCTGACGAGCATCAAAATAATCGACGTCAAGTCAGAGGTGGCGAA 1979
Db 1743 ATAGGCTCCGCGCCCTGACGAGCATCAAAATAATCGACGTCAAGTCAGAGGTGGCGAA 1684
Qy 1980 ACCGACAGGACTATAAGATACAGGGGTTTTCCCTCGGAAGCTCCCTCGTGGCTCTC 2039
Db 1683 ACCGACAGGACTATAAGATACAGGGGTTTTCCCTCGGAAGCTCCCTCGTGGCTCTC 1624
Qy 2040 CTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGG 2099
Db 1623 CTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGG 1564
Qy 2100 CGCTTTCTCATAGCTCAAGCTGTAGGTATCTCAGTTTCGGTGTAGGTCTGCTCCAAGC 2159
Db 1563 CGCTTTCTCATAGCTCAAGCTGTAGGTATCTCAGTTTCGGTGTAGGTCTGCTCCAAGC 1504
Qy 2160 TGGGCTGTGTGACGAACCCCGCTTCAGCCGACCGCTCGCCTTATCCGCTAACTATC 2219
Db 1503 TGGGCTGTGTGACGAACCCCGCTTCAGCCGACCGCTCGCCTTATCCGCTAACTATC 1444
Qy 2220 GTCTTGAGTCCAAACCGGTAAAGACACGACTTATCGCCACTGCGCAGCAGCCACTGCTGAACA 2279
Db 1443 GTCTTGAGTCCAAACCGGTAAAGACACGACTTATCGCCACTGCGCAGCAGCCACTGCTGAACA 1384
Qy 2280 GGAATTAGCAGCAGGATGATGATGCGGTGTCTACAGATTTCTTGAAGTGTGGCTAACT 2339
Db 1383 GGAATTAGCAGCAGGATGATGATGCGGTGTCTACAGATTTCTTGAAGTGTGGCTAACT 1324
Qy 2340 ACGGCTACACTAGAAAGACGATATTTGGTATCTCGCTCTGCTGAAGCAGTTACTCTCG 2399
Db 1323 ACGGCTACACTAGAAAGACGATATTTGGTATCTGCGCTCTGCTGAAGCAGTTACTCTCG 1264
Qy 2400 GAAAAAGAGTTGGTAGCTCTTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTT 2459
Db 1263 GAAAAAGAGTTGGTAGCTCTTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTT 1204
Qy 2460 TTGTTTGCAGCAGCAGATACGCGCAGAAAAAAGGATCTCAAGAGATCTCTTTGATCT 2519
Db 1203 TTGTTTGCAGCAGCAGATACGCGCAGAAAAAAGGATCTCAAGAGATCTCTTTGATCT 1144
Qy 2520 TTTCTACGGGCTGTGACGCTCAGTTGGACGAAACTCAGTTTAAAGGATTTTGGTCATGA 2579
Db 1143 TTTCTACGGGCTGTGACGCTCAGTTGGACGAAACTCAGTTTAAAGGATTTTGGTCATGA 1084
Qy 2580 GATTATCAAAAAGGATCTTCACTAGATCTCTTTAAATTAATAAATAAGTTTAAATCAA 2639
Db 1083 GATTATCAAAAAGGATCTTCACTAGATCTCTTTAAATTAATAAATAAGTTTAAATCAA 1024
Qy 2640 TCTAAAGTATATATAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCATGATGAGGAC 2699
Db 1023 TCTAAAGTATATATAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCATGATGAGGAC 964
Qy 2700 CTATCTCAGCGATCTGCTATTTCTGTTTCATCATAGTTGCTGACTCCCGCTCGTGTAGA 2759
Db 963 CTATCTCAGCGATCTGCTATTTCTGTTTCATCATAGTTGCTGACTCCCGCTCGTGTAGA 904

```
QY 2760 TAACTACGATACGGAGGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGGAGACC 2819
Db |||
QY 903 TAACTACGATACGGAGGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGGAGACC 844
Db |||
QY 2820 CACGCTCACGGCTCCAGATTTATCAGCAATAAACCCAGCAGCCGGAAGGGCCGAGCGCA 2879
Db |||
QY 843 CACGCTCACGGCTCCAGATTTATCAGCAATAAACCCAGCAGCCGGAAGGGCCGAGCGCA 784
Db |||
QY 2880 GAAGTGGTCTCAACTTTTATCGGCTCCATCCAGTCTTAATTTGTCGGGGAAGCTTA 2939
Db |||
QY 783 GAAGTGGTCTCAACTTTTATCGGCTCCATCCAGTCTTAATTTGTCGGGGAAGCTTA 724
Db |||
QY 2940 GAGTAAGTAGTTCGCCAGTAAATAGTTTCGGCAAGTTGTCGCAATGTCACAGCATCG 2999
Db |||
QY 723 GAGTAAGTAGTTCGCCAGTAAATAGTTTCGGCAAGTTGTCGCAATGTCACAGCATCG 664
Db |||
QY 3000 TGGTGTCCAGCTCGTCTGGTGTATGGCTTCATTAGCTCCGTTCCCAACCATCAAGC 3059
Db |||
QY 663 TGGTGTCCAGCTCGTCTGGTGTATGGCTTCATTAGCTCCGTTCCCAACCATCAAGC 604
Db |||
QY 3060 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAGCTCTTTCGGTCTCCGATCG 3119
Db |||
QY 603 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAGCTCTTTCGGTCTCCGATCG 544
Db |||
QY 3120 TTGTCAGAAAGTAGTTGGCCGAGTGTATCACTCATGTTATGCGAGCACTGCATAATT 3179
Db |||
QY 543 TTGTCAGAAAGTAGTTGGCCGAGTGTATCACTCATGTTATGCGAGCACTGCATAATT 484
Db |||
QY 3180 CTCCTACTGTGATGCCATCCGTAAGATGCTTTTCTGTGACTGTGAGTACTCAACCAAGT 3239
Db |||
QY 483 CTCCTACTGTGATGCCATCCGTAAGATGCTTTTCTGTGACTGTGAGTACTCAACCAAGT 424
Db |||
QY 3240 CATTTCTGAGAATAGTGTATGCGCGCAGCGAGTTGCTCTTTCGGCGGTCAATACCGGATA 3299
Db |||
QY 423 CATTTCTGAGAATAGTGTATGCGCGCAGCGAGTTGCTCTTTCGGCGGTCAATACCGGATA 364
Db |||
QY 3300 ATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTATCATTTGGAAACGTTCTTCGGGCG 3359
Db |||
QY 363 ATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTATCATTTGGAAACGTTCTTCGGGCG 304
Db |||
QY 3360 GAAACTCTCAAGGATCTTACCGCTGTTCAGATCCAGTTCGATGTAAACCACTCTGTGCAC 3419
Db |||
QY 303 GAAACTCTCAAGGATCTTACCGCTGTTCAGATCCAGTTCGATGTAAACCACTCTGTGCAC 244
Db |||
QY 3420 CCAACTGATCTTTCAGCATCTTTTACTTTTCCACAGCGTTTCTGGGTGAGCAAAAAACAGGAA 3479
Db |||
QY 243 CCAACTGATCTTTCAGCATCTTTTACTTTTCCACAGCGTTTCTGGGTGAGCAAAAAACAGGAA 184
Db |||
QY 3480 GGCAAAATGCCCAAAAAAGGAATAAGGGCGACACAGGAAATGTTGAATATCTCATCTCT 3539
Db |||
QY 183 GGCAAAATGCCCAAAAAAGGAATAAGGGCGACACAGGAAATGTTGAATATCTCATCTCT 124
Db |||
QY 3540 TCCTTTTTCATATTTATGAAGTATTTATCAGGTTATGTTCTCATGCGGATACATAT 3599
Db |||
QY 123 TCCTTTTTCATATTTATGAAGTATTTATCAGGTTATGTTCTCATGCGGATACATAT 64
Db |||
QY 3600 TTGAATGTTATTTAGAAAAATAAACAAATAGGGGTTCCCGCGCATTTTCCCGGAAAGTGC 3659
Db |||
QY 63 TTGAATGTTATTTAGAAAAATAAACAAATAGGGGTTCCCGCGCATTTTCCCGGAAAGTGC 4
Db |||
QY 3660 CAC 3662
Db |||
QY 3 CAC 1
```

RESULT 5
US-10-684-134-32/c
; Sequence 32, Application US/10684134
; Publication No. US20040142433A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5392-CIP

```
; CURRENT APPLICATION NUMBER: US/10/684,134  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: 60/402,342  
; PRIOR FILING DATE: 2002-08-08  
; PRIOR APPLICATION NUMBER: 10/066,390  
; PRIOR FILING DATE: 2002-02-01  
; PRIOR APPLICATION NUMBER: 60/268,785  
; PRIOR FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 60/266,386  
; PRIOR FILING DATE: 2002-02-02  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 32  
; LENGTH: 3637  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Encodes Cycle 3 BFP gene  
US-10-684-134-32  
  
Query Match 72.8%; Score 2665.4; DB 18; Length 3637;  
Best Local Similarity 83.7%; Pred. No. 0;  
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;  
  
QY 1 CTAATTTGTAACGGTTAATAATTTTGTAAAAATTCGGTTAAATTTTGTAAATCAGCTC 60  
Db |||  
QY 3637 CTAATTTGTAACGGTTAATAATTTTGTAAAAATTCGGTTAAATTTTGTAAATCAGCTC 3578  
Db |||  
QY 61 ATTTTTTAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 120  
Db |||  
QY 3577 ATTTTTTAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 3518  
Db |||  
QY 121 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGGACTC 180  
Db |||  
QY 3517 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGGACTC 3458  
Db |||  
QY 181 CAACTGTAAGGGCGCAAAACCGTCTATCAGGGCGATGCGCCACTACGTGAACCATCACC 240  
Db |||  
QY 3457 CAACTGTAAGGGCGCAAAACCGTCTATCAGGGCGATGCGCCACTACGTGAACCATCACC 3398  
Db |||  
QY 241 CTAATCAAGTTTTTTGGGTCGAGTCCGTAAGACACTAAATCGGAACCCCTAAAGGGAG 300  
Db |||  
QY 3397 CTAATCAAGTTTTTTGGGTCGAGTCCGTAAGACACTAAATCGGAACCCCTAAAGGGAG 3338  
Db |||  
QY 301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGCGCAACGTCGGGAGAAAGGAGGGAAGAA 360  
Db |||  
QY 3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGCGCAACGTCGGGAGAAAGGAGGGAAGAA 3278  
Db |||  
QY 361 AGCGAAAGAGCGGCGCTAGGGCGTGGCAAGTGTAGCGGTCACTGCTGGCGTAAACAC 420  
Db |||  
QY 3277 AGCGAAAGAGCGGCGCTAGGGCGTGGCAAGTGTAGCGGTCACTGCTGGCGTAAACAC 3218  
Db |||  
QY 421 CACACCGCGCGCTTAATATGCGCCCTACAGGGCGTCCCATTCGCCATTTCAGGCTGCG 480  
Db |||  
QY 3217 CACACCGCGCGCTTAATATGCGCCCTACAGGGCGTCCCATTCGCCATTTCAGGCTGCG 3158  
Db |||  
QY 481 CAACTGTTGGGAAGGGCGATCCGTCGGGCTCTTCGCTATTATACGCCAGTGGCGAAAGG 540  
Db |||  
QY 3157 CAACTGTTGGGAAGGGCGATCCGTCGGGCTCTTCGCTATTATACGCCAGTGGCGAAAGG 3098  
Db |||  
QY 541 GGGATGTGTCGAAGCGGATTAAGTTGGTAAACGCCAGGGTTTTTCCAGTCAGGAGTTG 600  
Db |||  
QY 3097 GGGATGTGTCGAAGCGGATTAAGTTGGTAAACGCCAGGGTTTTTCCAGTCAGGAGTTG 3038  
Db |||  
QY 601 TAAACGAGCGCGCAGTCAGCGCGCTCGTTTACAGCTATTTCAGTTTTTGAACCCGTCGAGGACGG 660  
Db |||  
QY 3037 TAAACGAGCGCGCAGTCAGCGCGCTCGTTTACAGCTATTTCAGTTTTTGAACCCGTCGAGGTCGA 2978  
Db |||  
QY 661 GCAGACTCGCGGTCAAAATGTTTATACAGCTGTATGAGGAGATGGAAGATGCTTCGACAC 720  
Db |||  
QY 2977 CCGCGGTGGCGCGCTCTAGAACTAGTGTGATCCCGCGGCTGCAGGAATTTCTTATTGT 2918  
Db |||  
QY 721 GCTGCAGAACACGCGAGCTAGATTAAACCCCTAGAAAAGATAATCATATTTGTGACGTACGTTAA 780
```

Db 2917 ATAGTTTCATCCATGTCATGTAATCCAGCAGCAGTTCACAACTCAAGAAGGACCATGT 2858
QY 781 AGATAATC-ATCGGTAAATATGACCATGGGATCTGTAATACGACTCACTATAGGGCGAA 839
Db 2857 GGTCAAGCTTTTCGTTGGGATCTTTTCGAAGGGCAGATTGTGTCAGAGGTAATGGTTGT 2798
QY 840 TTGGGTACCGGGCCCCCTCGAGGTGCAAGGTATCGGATAGCTTGAATGATATCGAATTCCTG 899
Db 2797 CTGGTAAAGGACAGGGCCATCGCCAAATTGGAGTATTTTGTGATAAGTGGTCTGTAGTT 2738
QY 900 CAGCCCGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGGGTGGAGCTCCAGCTTTTG 959
Db 2737 GAAAGGATCCATCTTCAATGTTGTGGCGAAATTTTGAAGTTAGCTTTGATTCATCTCTTT 2678
QY 960 TTCCCTTTAGTGAGGGTTAAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTA 1019
Db 2677 GTTTGTCTGCCGTGATATACATTTGTGTGATTAAAGTTGTACTCGAGTTTGTGTCCGA 2618
QY 1020 GGGTTAACTAGCTGCATCAGGATCATATCGTCGGGTCTTTTCCGGCTCAGTTCATCGC 1079
Db 2617 GAAATGTTCCATCTCTTTTAAATCAATACCTTTTAACTCGATAC-----GATTAAACAAG 2563
QY 1080 CCAAGCTGGCGTATCTGGGCATCGGGAGGAGAAAGCCCGTCTTTTCCGGGAGGTT 1139
Db 2562 GGTATCACCTTCAAACTTTGACTTCAGCACGGGCTCTGTAGTTCCCGTCACTCTTTGAAAGA 2503
QY 1140 GAAGCGGATGGAAGAGTTTCCCGAGGATGACTGCTGCTGCATTTGAGTTCAGCGAAAA 1199
Db 2502 TATAGTGGTTCCTGTACATAACCTTTCGGGATGGCACTCTTGAAAAAGTCAATGCCGTTT 2443
QY 1200 CGCAGTTTACCATGATGATTCGGGAAGTGTGGCCATGTCAGCGCTTTTAAACGGTGAAC 1259
Db 2442 CATATGATCCGGATAAACGAAAGACATTTGAACCATGAGAGAAAGTAGTGACAAAGTGT 2383
QY 1260 TTGCTTACGGCCACTGGGATACCAAGTTTCGTGGCGCTTTTCCGGACACAGTTCCGGATG 1319
Db 2382 TGGCCATGGAACAGGTAGTTTTCAGTAGTAGTGCMAATAAAATTAAGTGAAGCTTTCCGTA 2323
QY 1320 GTCAGCCGAGCGCATCAGNAACCGAACAATACCGGCGACAGCGGAATGCGTGCC 1379
Db 2322 TGTAGCATCACCTTCACCTCTCCACTCAGCAAAAAATTTGTGCCATTAACATCACCATC 2263
QY 1380 GGTGTGAGATTAATGACAGCGGTGCGCGCTGGGATATACGTCAAGCGAGGACGGGTAT 1439
Db 2262 TAAATCAACAAGAT-----TGGGACAACTCCAGTGAAAGTTC 2224
QY 1440 CTGCTGGATCCGCGAATAATGGACATGGATACCCCCGTGAGTTACCCGCGGGCGCGCT 1499
Db 2223 TTCTCCTTTACTCATCGGTACCCAGCTTTTGTTCCTTTTAGTGAGGGTTAAATTGCGCGCT 2164
QY 1500 TGGCGTAAATCATGTTATAGTGTTCCTGTGTGAATTTATTCGGTCAACATTCAC 1559
Db 2163 TGGCGTAAATCATGTTATAGTGTTCCTGTGTGAATTTATTCGGTCAACATTCAC 2104
QY 1560 ACAACATACGAGCGGGAAGCATAAAGTGTAAAGCTTGGGGTGCCTTAATGAGTGAGCTAAC 1619
Db 2103 ACAACATACGAGCGGGAAGCATAAAGTGTAAAGCTTGGGGTGCCTTAATGAGTGAGCTAAC 2044
QY 1620 TCACATTAATTTGCGTTGGCTCATCTGCCCGCTTTCCAGTCGGGAAAACCTGTGTCGCCAGC 1679
Db 2043 TCACATTAATTTGCGTTGGCTCACTGCCCGCTTTCCAGTCGGGAAAACCTGTGTCGCCAGC 1984
QY 1680 TGCATTAATGAATGGCCAAACCGCGGGAGAGGGGTTTCGGTATTTGGGCGCTCTTCGG 1739
Db 1983 TGCATTAATGAATGGCCAAACCGCGGGAGAGGGGTTTTCGGTATTTGGGCGCTCTTCGG 1924
QY 1740 CTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1799
Db 1923 CTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1864
QY 1800 ACTCAAAGCGGTAATACGGTTATCCACAGAAATACGGGATAACCGCAGGAAAGAACATGT 1859
|||||

Db 1863 ACTCAAAGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACGACGAAAAAGAACATGT 1804
QY 1860 GAGCAAAAGGCGCAGCAAAAGGCCAGGAACCGTTAAAAAAGCCGCGTTCGTGGCGTTTTTCC 1919
Db 1803 GAGCAAAAGGCGCAGCAAAAGGCCAGGAACCGTTAAAAAAGCCGCGTTCGTGGCGTTTTTCC 1744
QY 1920 ATAGGCTCCGCGCCCCCTCAGCAGCATCAAAAAATCGACGCTCAAGTCAAGAGTGGCGAA 1979
Db 1743 ATAGGCTCCGCGCCCCCTCAGCAGCATCAAAAAATCGACGCTCAAGTCAAGAGTGGCGAA 1684
QY 1980 ACCGACAGGACTATAAAGATACAGAGCGTTTCCCTCGGAAGCTCCCTCGTGCCTCTC 2039
Db 1683 ACCGACAGGACTATAAAGATACAGAGCGTTTCCCTCGGAAGCTCCCTCGTGCCTCTC 1624
QY 2040 CTGTTCCGACCCCTCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAACGCTGG 2099
Db 1623 CTGTTCCGACCCCTCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAACGCTGG 1564
QY 2100 CGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGGTGTAGTCTCGTTCGCTCAAAGC 2159
Db 1563 CGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGGTGTAGTCTCGTTCGCTCAAAGC 1504
QY 2160 TGGGCTGTGTCGACGAAACCCCGTTCAGCCCGACCGCTGCGCTTATCCGTTAACTATC 2219
Db 1503 TGGGCTGTGTCGACGAAACCCCGTTCAGCCCGACCGCTGCGCTTATCCGTTAACTATC 1444
QY 2220 GTCTTGAGTCCAAACCCGTTAAGACACAGCTTATCGCCACTGGCAGCAGCCTGTTAAACA 2279
Db 1443 GTCTTGAGTCCAAACCCGTTAAGACACAGCTTATCGCCACTGGCAGCAGCCTGTTAAACA 1384
QY 2280 GGATTAGCAGACGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCCTAACT 2339
Db 1383 GGATTAGCAGACGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCCTAACT 1324
QY 2340 ACGGTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACTTCG 2399
Db 1323 ACGGTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACTTCG 1264
QY 2400 GAAAAAGATTGGTAGCTCTTTGATCCGGCAAAACAAACACCGCTGCTGAGCGTGGTTTTT 2459
Db 1263 GAAAAAGATTGGTAGCTCTTTGATCCGGCAAAACAAACACCGCTGCTGAGCGTGGTTTTT 1204
QY 2460 TTGTTTGAACGACGAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCTTTTGATCT 2519
Db 1203 TTGTTTGAACGACGAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCTTTTGATCT 1144
QY 2520 TTTCTACGGGCTCTGACGCTCAGTGGACGAAAACTCAGTTTAAGGATTTTGGTCATGA 2579
Db 1143 TTTCTACGGGCTCTGACGCTCAGTGGACGAAAACTCAGTTTAAGGATTTTGGTCATGA 1084
QY 2580 GATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCAA 2639
Db 1083 GATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCAA 1024
QY 2640 TCTAAAGTATATATGAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTAGGCGAC 2699
Db 1023 TCTAAAGTATATATGAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTAGGCGAC 964
QY 2700 CTATCTCAGCGATCTGTCTATTTGTTTCATCCATAGTTGCGCTGACTCCCGCTGTTAGA 2759
Db 963 CTATCTCAGCGATCTGTCTATTTGTTTCATCCATAGTTGCGCTGACTCCCGCTGTTAGA 904
QY 2760 TAACTACGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCGAGACC 2819
Db 903 TAACTACGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCGAGACC 844
QY 2820 CACGCTCAACCGGCTCCAGATTTATCAGCAATTAACAGCCAGCGGAAGGGCCGAGCGCA 2879
Db 843 CACGCTCAACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCGGAAGGGCCGAGCGCA 784
QY 2880 GAAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTTAATTTGTTGGCGGAGCTA 2939
Db 783 GAAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTTAATTTGTTGGCGGAGCTA 724

```
QY 2940 GAGTAAGTAGTTCGCCAGTAAATAGTTTGGCGAACGTTGTTGCCATGTTGCTACAGGCATCG 2999
Db 723 GAGTAAGTAGTTCGCCAGTAAATAGTTTGGCGAACGTTGTTGCCATGTTGCTACAGGCATCG 664
QY 3000 TGGTGTCCAGCTCGTCTGTTGGTATCGCTTCATTCAGCTCCGTTCCCAACGATCAAGGC 3059
Db 663 TGGTGTCCAGCTCGTCTGTTGGTATCGCTTCATTCAGCTCCGTTCCCAACGATCAAGGC 604
QY 3060 GAGTTACATGATCCCCCATGTTGTGCAAAAAGCGTTAGCTCTTCGGTCCCTCCGATCG 3119
Db 603 GAGTTACATGATCCCCCATGTTGTGCAAAAAGCGTTAGCTCTTCGGTCCCTCCGATCG 544
QY 3120 TTGTCAGAGTAAGTTGGCGCAGTGTATCACTCATGTTATGTCAGGAGCACTGCATAAAT 3179
Db 543 TTGTCAGAGTAAGTTGGCGCAGTGTATCACTCATGTTATGTCAGGAGCACTGCATAAAT 484
QY 3180 CTCCTTACTCTGATGCCATCCGTAAGATGCTTTTCTGTGACTGCTGAGTACTCAACCAAGT 3239
Db 483 CTCCTTACTCTGATGCCATCCGTAAGATGCTTTTCTGTGACTGCTGAGTACTCAACCAAGT 424
QY 3240 CATCTGAGATAGTGTATGCGCGCAGCGAGTTGCTCTTGGCCCGGCTCAATACGGGATA 3299
Db 423 CATCTGAGATAGTGTATGCGCGCAGCGAGTTGCTCTTGGCCCGGCTCAATACGGGATA 364
QY 3300 ATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCATTTGGAAACGTTCTTCGGGCG 3359
Db 363 ATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCATTTGGAAACGTTCTTCGGGCG 304
QY 3360 GAAACTCTCAGGATCTTACCGCTGTGAGATCCAGTTGCTCTTGGCCCGGCTCAATACGGGATA 3419
Db 303 GAAACTCTCAGGATCTTACCGCTGTGAGATCCAGTTGCTCTTGGCCCGGCTCAATACGGGATA 244
QY 3420 CCAACTGATCTTCAGCATCTTTTACTTTTCCACAGCGTTTCTGGTGGAGCAAAACAGGAA 3479
Db 243 CCAACTGATCTTCAGCATCTTTTACTTTTCCACAGCGTTTCTGGTGGAGCAAAACAGGAA 184
QY 3480 GGCATAATCCGCAAAAAGGGAATAAGGCGCACGGAATGTTGAATCTCATCTCT 3539
Db 183 GGCATAATCCGCAAAAAGGGAATAAGGCGCACGGAATGTTGAATCTCATCTCT 124
QY 3540 TCCCTTTTCAATATTTAAGCATTTATCAGGTTTATGCTCATGAGGAGATACATAT 3599
Db 123 TCCCTTTTCAATATTTAAGCATTTATCAGGTTTATGCTCATGAGGAGATACATAT 64
QY 3600 TTGAATGTTATTTAGAAAAATAAACAATAGGGGTTCCGCGCATTTTCCCGAAAAAGTGC 3659
Db 63 TTGAATGTTATTTAGAAAAATAAACAATAGGGGTTCCGCGCATTTTCCCGAAAAAGTGC 4
QY 3660 CAC 3662
Db 3 CAC 1
```

RESULT 6

```
US-10-637-758-32/c
; Sequence 32, Application US/10637758
; Publication No. US20040180352A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/637, 758
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: US/10/280, 913
; PRIOR FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 60/402, 342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066, 390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268, 785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266, 386
```

```
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Encodes Cycle 3 BFP gene
US-10-637-758-32
```

```
Query Match 72.8%; Score 2665.4; DB 18; Length 3637;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;

QY 1 CTAAATGTTAAGCGTTAAATATTTTGTAAAAATTCGCGTTAAAAATTTTGTAAAAATTCAGCTC 60
Db 3637 CTAAATGTTAAGCGTTAAATATTTTGTAAAAATTCGCGTTAAAAATTTTGTAAAAATTCAGCTC 3578
QY 61 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 120
Db 3577 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 3518
QY 121 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTGGACTC 180
Db 3517 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTGGACTC 3458
QY 181 CAAAGTCAAAGGCGGCAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
Db 3457 CAAAGTCAAAGGCGGCAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398
QY 241 CTAATCAAGTTTTTTGGGTCGAGGTGCGCTAAAGCACTAAATTCGGAACCTTAAAGGAG 300
Db 3397 CTAATCAAGTTTTTTGGGTCGAGGTGCGCTAAAGCACTAAATTCGGAACCTTAAAGGAG 3338
QY 301 CCCCAGATTTAGACTTTGACGGGAAAGCCGGCAACGTTGGCGAGAAAGGAGGAGAA 360
Db 3337 CCCCAGATTTAGACTTTGACGGGAAAGCCGGCAACGTTGGCGAGAAAGGAGGAGAA 3278
QY 361 AGCGAAAGGAGCGGGCGCTAGGCGCTGCAAGTGTAGCGGTCACTGCGCGCTAACCCAC 420
Db 3277 AGCGAAAGGAGCGGGCGCTAGGCGCTGCAAGTGTAGCGGTCACTGCGCGCTAACCCAC 3218
QY 421 CACACCCCGCGCTTAAATGCGCGCTACAGGCGCGTCCCATTCGCCATTCAGGTGCG 480
Db 3217 CACACCCCGCGCTTAAATGCGCGCTACAGGCGCGTCCCATTCGCCATTCAGGTGCG 3158
QY 481 CAACTGTTGGAAAGGCGGATCGGTGCGGGCTCTTCGCTATTACGCCAGCTGCGGAAG 540
Db 3157 CAACTGTTGGAAAGGCGGATCGGTGCGGGCTCTTCGCTATTACGCCAGCTGCGGAAG 3098
QY 541 GGGATGCTGCAAGCGGATTAAGTTGGGTAAAGCCAGAGGTTTCCAGTCACGAGCTTG 600
Db 3097 GGGATGCTGCAAGCGGATTAAGTTGGGTAAAGCCAGAGGTTTCCAGTCACGAGCTTG 3038
QY 601 TAAACGACGCGCAGTGAAGCGCGCTCTGTTTCACTTCATTCGTTTTTGAACCCCGTGGAGACGG 660
Db 3037 TAAACGACGCGCAGTGAAGCGCGCTCTGTTTCACTTCATTCGTTTTTGAACCCCGTGGAGTCCA 2978
QY 661 GCAGACTCCGGTGCATAATGTTTACAGCGTGTAGGAGCAGATGAAGATGCTCGACAC 720
Db 2977 CCGCGTGGCGCGCTCTAGAACTAGTGGATCCCGCGGCTGCGAGGAATTTCTTATTTGT 2918
QY 721 GCTGCAGAACACGCGAGCTAGATTAAACCTAGAAAAGATAATCATATTGTGACGTCAGTTAA 780
Db 2917 ATAGTTTATCCATGCCATGTGTAAATCCAGCAGCAGTTACAAACTCAAGAAAGACCATGT 2858
QY 781 AGATAATC-ATGCGTAAAAATTCAGCGATGGGATCTGTAATACGACTCACTATATAGGCGAA 839
Db 2857 GGTCAAGCTTTTGGTGGGATCTTTTCAAGAAAGGCGAGATTTGTGTCGACAGGTAATGTTGT 2798
QY 840 TTGGGTACCGGGCCCCCTCTCGAGGTCGACGGTATCGAATGCTTGAATTCGAATTCCTG 899
Db 840 TTGGGTACCGGGCCCCCTCTCGAGGTCGACGGTATCGAATGCTTGAATTCGAATTCCTG 899
```

Db 2797 CTGGTAAAGAGCAGGGCCCATCCCAAATGGAGTATTTTGTGATAATGGTCTGCTAGTT 2738
Qy 900 CAGCCGGGGATCCACATAGTTCTAGAGCGCGCCACCGCGGTGGAGCTCCAGCTTTTG 959
Db 2737 GAACGGATCCATCTTCAATGTTGTGGGAAATTTGAAAGTTAGCTTTGATTCATTTCTTT 2678
Qy 960 TTCCCTTTAGTGAGGGTTAAATTTAGATCCCATCGTCAATTTTTPACGACACTATCTTTCTA 1019
Db 2677 GTTTGCTGCGGTGATATACATTTGTGTGAGTTTAAAGTTGTACTCGAGTTTGTGTCCGA 2618
Qy 1020 GGGTTAACTAGCTGCATCAGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCAATCGC 1079
Db 2637 GAATGTTTCCATCTTTCTTTAAATCAATAACCTTTTAACTCGATAC-----GATTAAACA 2563
Qy 1080 CCAAGCTGCGCTATCTGGGCATCCGGGAGGAAGCCGTCCTTTTCCCGCAGGTT 1139
Db 2562 GGTATCACCTTCAAACTTTGACTTCAGCAGCGGTCTTTAGTTCCTCGTCACTCTTTGAAAGA 2503
Qy 1140 GAAGCGCATGGAAGAGTTTTCGAGGATGACTCTGCTGCAATTGACGTTTGAGCGAAAA 1199
Db 2502 TATAGTGGTTCCTGTACATAACCTTCGGGCATGGCACTCTTTGAAAAAGTCATGCCGTTT 2443
Qy 1200 CGCAGTTTACCATGATGATTCGGGAAGGTGGGCCATGCAGCCCTTTAAACGGTGAACGT 1259
Db 2442 CATATGATCCGGATAACGAGAAAAAGCATTTGAACACCATGAGAGAAAGTAGTGACAAAGTGT 2383
Qy 1260 TTGCTTCAAGGCACCTGGGATACAGATTTCGTGCGGGCTTTTCCGGACACAGTTCCGGATG 1319
Db 2382 TGGCATGGAACAGGTAGTTTTCAGTAGTAGTGCAATAATTTAAGTGTAAAGCTTTCCGTA 2323
Qy 1320 GTCCGCGGAAGCGCATCAGCAACCCGAACATACCGCGCAGCAGCGGAACCTGCCGTGCC 1379
Db 2322 TGAGCATCACCTTCACCTCTCCACTGACAGAAAAATTTGTGCCCAATTTAAACATCACCATC 2263
Qy 1380 GGTGTGAGATTAATGACAGCGGTGCGGGCTGGGATATAGTCAAGAGGAGCGGTAT 1439
Db 2262 TAAATTCACAAAGAAAT-----TGGGCAACTCCAGTGAAGAAAGTTC 2224
Qy 1440 CTGGCTGGATGCCGAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGCGCT 1499
Db 2223 TTCTCCTTTACTCATCGGTACCCAGCTTTGTTCCTTTTAGTGGGGTTAATTGGCGCT 2164
Qy 1500 TGGCGTAAATCATGTGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCAAAATTCAC 1559
Db 2163 TGGCGTAAATCATGTGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCAAAATTCAC 2104
Qy 1560 ACAACATACGAGCGCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAATGAGTGAGCTAAC 1619
Db 2103 ACAACATACGAGCGCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAATGAGTGAGCTAAC 2044
Qy 1620 TCACATTAATTTGCGTTGCGCTCACTGCCGCTTTCCAGTCCGGAACCTGTCGTGCCAGC 1679
Db 2043 TCACATTAATTTGCGTTGCGCTCACTGCCGCTTTCCAGTCCGGAACCTGTCGTGCCAGC 1984
Qy 1680 TGCATTAATGAAATCGGCCAAACCGCGGGGAGAGCGGTTTGGGTATTTGGGGCTCTTCGG 1739
Db 1983 TGCATTAATGAAATCGGCCAAACCGCGGGGAGAGCGGTTTGGGTATTTGGGCGCTCTTCGG 1924
Qy 1740 CTTCTCGCTCACTGACTCGCTCGCTCGGTGTTGCGGTCGGGCGAGCGGTATCAGCTC 1799
Db 1923 CTTCTCGCTCACTGACTCGCTCGCTCGGTGTTGCGGTCGGGCGAGCGGTATCAGCTC 1864
Qy 1800 ACTCAAGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACGACGAAAGAACATGT 1859
Db 1863 ACTCAAGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACGACGAAAGAACATGT 1804
Qy 1860 GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGGTAAAGCCCGGTTGCTGGCGTTTTTTC 1919
Db 1803 GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGGTAAAGAGGCCCGGTTGCTGGCGTTTTTTC 1744
Qy 1920 ATAGGCTCGGCCCTCGACGAGCATCAAAATAATCGACGCTCAAGTCAGAGGTGGCGAA 1979
Db 1743 ATAGGCTCGGCCCTCGACGAGCATCAAAATAATCGACGCTCAAGTCAGAGGTGGCGAA 1684

Qy 1980 ACCGACAGGACTATAAGATACCAAGCGTTTCCCTCGGAAGCTCCCTCGTGGCTCTC 2039
Db 1683 ACCGACAGGACTATAAGATACCAAGCGTTTCCCTCGGAAGCTCCCTCGTGGCTCTC 1624
Qy 2040 CTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTCGGGAAGCGTGG 2099
Db 1623 CTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTCGGGAAGCGTGG 1564
Qy 2100 CGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTAGGTGCTTCCCTCCAAGC 2159
Db 1563 CGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTAGGTGCTTCCCTCCAAGC 1504
Qy 2160 TGGGCTGTGTGCAGAAACCCCGCTTACGCCGACCGCTGCGCTTATTCGGTAACTATC 2219
Db 1503 TGGGCTGTGTGCAGAAACCCCGCTTACGCCGACCGCTGCGCTTATTCGGTAACTATC 1444
Qy 2220 GTCTTGAAGTCCAAACCCCGTAAAGACAGACTTATGCCACTGCGCAGCAGCACTGGTAACA 2279
Db 1443 GTCTTGAAGTCCAAACCCCGTAAAGACAGACTTATGCCACTGCGCAGCAGCACTGGTAACA 1384
Qy 2280 GGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT 2339
Db 1383 GGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT 1324
Qy 2340 ACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACTTTCG 2399
Db 1323 ACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACTTTCG 1264
Qy 2400 GAAAAAGATTGGTAGTCTTTGATCCCGCAACAAACCAACCGCTGGTAGCGGTGTTTTT 2459
Db 1263 GAAAAAGATTGGTAGTCTTTGATCCCGCAACAAACCAACCGCTGGTAGCGGTGTTTTT 1204
Qy 2460 TTGTTTGCAGCAGCAGATTACGGCAGAAAAAAGGATCTCAGAAAGTCTTTTGATCT 2519
Db 1203 TTGTTTGCAGCAGCAGATTACGGCAGAAAAAAGGATCTCAGAAAGTCTTTTGATCT 1144
Qy 2520 TTCTTACGGGCTCTGACGCTCAGTGGAAACGAAACTCAGTTTAAAGGATTTTGGTCATGA 2579
Db 1143 TTTCTACGGGCTCTGACGCTCAGTGGAAACGAAACTCAGTTTAAAGGATTTTGGTCATGA 1084
Qy 2580 GATTATCAAAAAGGATCTTACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAA 2639
Db 1083 GATTATCAAAAAGGATCTTACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAA 1024
Qy 2640 TCTAAAGTATATACGATTAACCTTGGTCTGACAGTTACCAATGCTTAAATCAGTAGGAC 2699
Db 1023 TCTAAAGTATATACGATTAACCTTGGTCTGACAGTTACCAATGCTTAAATCAGTAGGAC 964
Qy 2700 CTATCTCAGCGATCTGTCTATTTTGGTTCATCATAGTTTGCCTGACTCCCGCTGCTGTAGA 2759
Db 963 CTATCTCAGCGATCTGTCTATTTTGGTTCATCATAGTTTGCCTGACTCCCGCTGCTGTAGA 904
Qy 2760 TAACTACGATACGGAGGGCTTACCATCTGCCCCAGTGTGCAATGATACCGCAGACC 2819
Db 903 TAACTACGATACGGAGGGCTTACCATCTGCCCCAGTGTGCAATGATACCGCAGACC 844
Qy 2820 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCCAGCCAGCGGAGGCGGAGCGCA 2879
Db 843 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCCAGCCAGCGGAGGCGGAGCGCA 784
Qy 2880 GAAAGTGGTCTGCAACTTTATCTCCGCTCCATCCAGTCTATTAAATTTGTTGCCGGGAAGCTA 2939
Db 783 GAAAGTGGTCTGCAACTTTATCTCCGCTCCATCCAGTCTATTAAATTTGTTGCCGGGAAGCTA 724
Qy 2940 GAGTAAAGTAGTTCCGCAAGTTAATAGTTTGGCGAAGCTTTGTTGCCATTTGCTACAGGCATCG 2999
Db 723 GAGTAAAGTAGTTCCGCAAGTTAATAGTTTGGCGAAGCTTTGTTGCCATTTGCTACAGGCATCG 664
Qy 3000 TGGTGTCAAGCTCGCTGTTTGGTATGGCTTCAATTCAGCTCCGTTTCCCAACGATCAAGC 3059
Db 663 TGGTGTCAAGCTCGCTGTTTGGTATGGCTTCAATTCAGCTCCGTTTCCCAACGATCAAGC 604

```
QY 3060 GAGTTACATGATCCCCCATGTTGTGCAGAAAGCGGTTAGCTCCTTCGGTCTCCGATCG 3119
DB 603 GAGTTACATGATCCCCCATGTTGTGCAGAAAGCGGTTAGCTCCTTCGGTCTCCGATCG 544
QY 3120 TTGTGAGAGTAAGTTGGCGGAGTGTATCACTCATGTGTATGGCAGCACTGCATAAT 3179
DB 543 TTGTGAGAGTAAGTTGGCGGAGTGTATCACTCATGTGTATGGCAGCACTGCATAAT 484
QY 3180 CTCCTACTGTCATGCCATCCGTAAGATGTTTTCTGTGACTGTTGAGTACTCAACCACT 3239
DB 483 CTCCTACTGTCATGCCATCCGTAAGATGTTTTCTGTGACTGTTGAGTACTCAACCACT 424
QY 3240 CATTCTGAGATAGTGTATGCCGCGACGAGTTGCTCTTCCCGCGGTCAATACGGGATA 3299
DB 423 CATTCTGAGATAGTGTATGCCGCGACGAGTTGCTCTTCCCGCGGTCAATACGGGATA 364
QY 3300 ATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAACGTTCTTCGGGCG 3359
DB 363 ATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAACGTTCTTCGGGCG 304
QY 3360 GAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGCAC 3419
DB 303 GAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGCAC 244
QY 3420 CCAACTGATCTTCAGCATCTTTTACTTTTCCAGCGGTTTCTGGTGAGCAAAAACAGGAA 3479
DB 243 CCAACTGATCTTCAGCATCTTTTACTTTTCCAGCGGTTTCTGGTGAGCAAAAACAGGAA 184
QY 3480 GGCAGAAATGCCCAAAAGGAAATAAGGCGGACACGGAATAATGTTGAATCTCATCTCT 3539
DB 183 GGCAGAAATGCCCAAAAGGAAATAAGGCGGACACGGAATAATGTTGAATCTCATCTCT 124
QY 3540 TCCCTTTTCAATATTTAAGCATTTATCAGGTTTATGTTCTCATGAGCGGATACATAT 3599
DB 123 TCCCTTTTCAATATTTAAGCATTTATCAGGTTTATGTTCTCATGAGCGGATACATAT 64
QY 3600 TTGAATGATTTAGAAAAATAAATAAGGTTTCCGCGCACATTTCCCGAAAGTGC 3659
DB 63 TTGAATGATTTAGAAAAATAAATAAGGTTTCCGCGCACATTTCCCGAAAGTGC 4
QY 3660 CAC 3662
DB 3 CAC 1
```

RESULT 7

```
US-10-066-390-4/c
; Sequence 4, Application US/10066390
; Publication NO. US20020146732A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Lindbo, John A.
; APPLICANT: Fitzmaurice, Wayne P.
; TITLE OF INVENTION: A Method of Increasing Complementarity
; TITLE OF INVENTION: In A Heteroduplex
; FILE REFERENCE: P-LG 4878
; CURRENT APPLICATION NUMBER: US/10/066,390
; CURRENT FILING DATE: 2002-02-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-066-390-4
```

```
Query Match 72.8%; Score 2664.8; DB 13; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;
```

```
QY 1 CTAATTTGTAAGCGTTAAATATTTTGTAAAAATTCGGGTTAAATTTTGTAAATACAGCTC 60
```

```
DB 3637 CTAATTTGTAAGCGTTAAATATTTTGTAAAAATTCGGGTTAAATTTTGTAAATACAGCTC 3578
QY 61 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTAAATCAAAAGAAATAGACCGA 120
DB 3577 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTAAATCAAAAGAAATAGACCGA 3518
QY 121 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTGACTC 180
DB 3517 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTGACTC 3458
QY 181 CAACGTCAAAGCGCGAAACCGTCTATCAGGCGCATGCGCCACTACGTGAACCATCACC 240
DB 3457 CAACGTCAAAGCGCGAAACCGTCTATCAGGCGCATGCGCCACTACGTGAACCATCACC 3398
QY 241 CTAATCAAGTTTTTGGGTCGAGTGCCTGAAAGCACTAAATCGGAACCCCTAAAGGGAG 300
DB 3397 CTAATCAAGTTTTTGGGTCGAGTGCCTGAAAGCACTAAATCGGAACCCCTAAAGGGAG 3338
QY 301 CCCCCGATTTAGACTTTGACGGGGAAAGCGCGGAAACGTGGCGAGAAAGGAGGAAGAA 360
DB 3337 CCCCCGATTTAGACTTTGACGGGGAAAGCGCGGAAACGTGGCGAGAAAGGAGGAAGAA 3278
QY 361 AGCGAAAGAGCGGCGCTAGGCGCTGCAAGTGTAGCGCTCACGCTGCGCTAACCCAC 420
DB 3277 AGCGAAAGAGCGGCGCTAGGCGCTGCAAGTGTAGCGCTCACGCTGCGCTAACCCAC 3218
QY 421 CACACCCGCGCGCTTAAATGCGCCGCTACAGGCGCGCTCCCATTCGCCATTCAGGCTGCG 480
DB 3217 CACACCCGCGCGCTTAAATGCGCCGCTACAGGCGCGCTCCCATTCGCCATTCAGGCTGCG 3158
QY 481 CAACTGTTGGGAAAGGCGGATCGGTGCGGCTCTTTCGCTATTACGCCAGTGCAGGAAGG 540
DB 3157 CAACTGTTGGGAAAGGCGGATCGGTGCGGCTCTTTCGCTATTACGCCAGTGCAGGAAGG 3098
QY 541 GGGATGCTGCTCAAGCGGATTAAGTTGGGTAAAGTGGGTAAAGTGGGTAAAGTGGGTAAAG 600
DB 3097 GGGATGCTGCTCAAGCGGATTAAGTTGGGTAAAGTGGGTAAAGTGGGTAAAGTGGGTAAAG 3038
QY 601 TAAAAACGAGCGGCGAGTGAAGCGGCTCGTTTCATTCACGTTTTTGAACCCGCTGAGGACGG 660
DB 3037 TAAAAACGAGCGGCGAGTGAAGCGGCTCGTTTCATTCACGTTTTTGAACCCGCTGAGGACGG 2978
QY 661 GCAGACTCGGTCGCAATGTGTTTACAGCGTGTAGGAGCAGATGAAGATGTCGACAC 720
DB 2977 CCGCGTGGCGG-----CCGCTCTAGAACTAGTGGATCCCC--CGG 2939
QY 721 GCTGCAGAACGCGAGCTAGATTAAACCTAGAAAGATATCATATTGTGACGTACGTTAA 780
DB 2938 GCTGCAGAACGCGAGCTAGATTAAACCTAGAAAGATATCATATTGTGACGTACGTTAA 2879
QY 781 AGATAATCATGCGTAAAAATTGACCGATGGGATCTGTAATACGACTCACTATAGGCGGAAT 840
DB 2878 CAAACTCAAGAAGGACCATGTGTCAGCGCTTTTCGTTGGGATCTTTCGAAAGGCGAGATT 2819
QY 841 TGGGTACCGGGCCCCCTCGAGGTCGACGGTATCGATGAAGTGTGATATCGAATTCCTGC 900
DB 2818 GTGTGCACAGGTAATGTTGTCTGTTAAAGGACAGGCGCATCGCCAAATTGGAGTATTT 2759
QY 901 AGCCCCGGGGATCCACTAGTCTTAGAGCGCGCCACCGCGGTGGAGCTCCAGCTTTTGT 960
DB 2758 GTTGATAATGTTCTGCTAGTTGAACGATCCATCTTCAATGTTGTGGCGAATTTGAAGT 2699
QY 961 TCCCTTTAGTGAGGTTAAATTAGATCCCATGGTCAATTTTACGACAGACTATCTTTCTAG 1020
DB 2698 TAGCTTTGATCCATCTTTTGTCTGTCGCGTGTATGATATACATTTGTGTGAG-TTATAG 2640
QY 1021 GGTAAATCTAGCTGCATCAGGATCATATCGTCGGGCTCTTTTTCGGCTCAGTCATCGCC 1080
DB 2639 TTGTACTCGAGTTGTGTCGAGAAATGTTTCCATCTTCTTTAAATCAATACCTTTTAAAC 2580
QY 1081 CAACTGCGGCTATCTGGGCATCGGGAGGAAAGACCGCTGCTTTTCCCGGAGTTG 1140
```


Db 2579 TCGATACGATTAAACAAGGGTATCACCTTCAAACCTTGACTTCAGCACGGGTCTGTAGTTC 2520
Qy 1141 AAGCGGCATGGAAGAGATTGCGGAGGATGACTGCTGCTGCATTCACGCTTGAGCGAAAC 1200
Db 2519 CGGTCACTTTGAAAGATATAGTGGTTCCTGTGATATAAACCTTCGGGCATGGCACTCTTG 2460
Qy 1201 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGACCGCTTTAAACGGTGAACCTGT 1260
Db 2459 AAAAAGTCATCGCGTTTCATATGATCCGGATAACGGGAAGACATTTGAACACCATTAAG 2400
Qy 1261 TCGTTACGCCACCTGGGATACAGTTGTCGTCGGCTTTTCGGGACACAGTTCCGGATGG 1320
Db 2399 AAAGTAGTGACAAGTGTGTGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAAATAAATTA 2340
Qy 1321 TCAGCCCAAGCGCATACGACCCGAAACAATACCGGCGACAGCCGGAACTGCCGTGCG 1380
Db 2339 AGGGTAAAGCTTTCGGTATGTAGCATCACCTTCACCTCTCCACTGACAGAAATTTGTGC 2280
Qy 1381 GTGTGCAGATTAATGACAGCGGTGCGGCTGGGATATTAAGTCAGCGAGGACGGGTATC 1440
Db 2279 CCAITTAACATCACCATCTAATTCAACAAGAAATGGGACAACTCCAGTGAAAAGTTCTTCT 2220
Qy 1441 CTGGCTGGATCCGCAGAAATGGAATGGAATACCCCTGGAGTTACCCGGCGGGCGCTT 1500
Db 2219 C---CTTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTAAATTTGGCGCGCTT 2163
Qy 1501 GCGTAAATCATGTCATAGCTGTTTCCTGTGTGAATTTGTTATCCGCTCACAATTCACA 1560
Db 2162 GCGTAAATCATGTCATAGCTGTTTCCTGTGTGAATTTGTTATCCGCTCACAATTCACA 2103
Qy 1561 CAACATACGAGCGGAAGCAATAAGTGTAAAGCTTGGGGTGCTTAATGAGTGAGCTAACT 1620
Db 2102 CAACATACGAGCGGAAGCAATAAGTGTAAAGCTTGGGGTGCTTAATGAGTGAGCTAACT 2043
Qy 1621 CACATTAATGCGTTGCGCTCATCTGCCGCTTTCAGTTCGGGAACCTGTGTGCGCAGCT 1680
Db 2042 CACATTAATGCGTTGCGCTCACTTGCCGCTTTCAGTTCGGGAACCTGTGTGCGCAGCT 1983
Qy 1681 GCATTAATGAATCGSCAAACGCGGGGAGAGGGGTTTGGTATTTGGCGCTCTTCGCG 1740
Db 1982 GCATTAATGAATCGGCCAAACGCGGGGAGAGCGGTTTGGTATTTGGCGCTCTTCGCG 1923
Qy 1741 TTCTCTCGTCACTGACTCGCTCGCTCGGTGCTTGGCTGGGCGAGCGGTATCAGCTCA 1800
Db 1922 TTCTCTCGTCACTGACTCGCTCGCTCGGTGCTTGGCTGGGCGAGCGGTATCAGCTCA 1863
Qy 1801 CTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATTAACGAGGAAAGACATGTG 1860
Db 1862 CTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATTAACGAGGAAAGACATGTG 1803
Qy 1861 AGCAAAAGCCAGCAAAAGCCAGGAACCGTAAGGCGCGTGTGCGGCTTTTTCCTCA 1920
Db 1802 AGCAAAAGCCAGCAAAAGCCAGGACCGTAAGGCGCGTGTGCGGCTTTTTCCTCA 1743
Qy 1921 TAGGCTCGCGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGCGGAAA 1980
Db 1742 TAGGCTCGCGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGCGGAAA 1683
Qy 1981 CCGACAGGACTATMAAGATACAGGCGGTTTCCCGCTGGAAAGTCCCTCGTGGCTCTCC 2040
Db 1682 CCGACAGGACTATMAAGATACAGGCGGTTTCCCGCTGGAAAGTCCCTCGTGGCTCTCC 1623
Qy 2041 TGTTCCGACCTGCGCTTACCGGATACCTGTGCGCCCTTCTCCCTTCGGAAGCGTGGC 2100
Db 1622 TGTTCCGACCTGCGCTTACCGGATACCTGTGCGCCCTTCTCCCTTCGGAAGCGTGGC 1563
Qy 2101 GCTTTCTCATAGCTCACGCTGTAGTATCTCAGTTTCGGTGTAGGTGCTTCGCTCCAAGCT 2160
Db 1562 GCTTTCTCATAGCTCACGCTGTAGTATCTCAGTTTCGGTGTAGGTGCTTCGCTCCAAGCT 1503
Qy 2161 GGGCTGTGTGACGAAACCCCGGTTTCAGCCGACCGCTGCGGCTTATCCGGTAACTATCG 2220
Db 1502 GGGCTGTGTGACGAAACCCCGGTTTCAGCCGACCGCTGCGGCTTATCCGGTAACTATCG 1443

Qy 2221 TCTTGAGTCCAAACCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAG 2280
Db 1442 TCTTGAGTCCAAACCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAG 1383
Qy 2281 GATTAGCAGACGAGGTATGTAGGCGGTCTACAGAGTTCTTGAAAGTGGTGGCTTAACTA 2340
Db 1382 GATTAGCAGACGAGGTATGTAGGCGGTCTACAGAGTTCTTGAAAGTGGTGGCTTAACTA 1323
Qy 2341 CGGCTACACTAAGAGGACAGTATTGGTATCTGGCTCTGTGAGCCAGTACCTTCGG 2400
Db 1322 CGGCTACACTAAGAGGACAGTATTGGTATCTGGCTCTGTGAGCCAGTACCTTCGG 1263
Qy 2401 AAAAAGAGTTCGTAGTCTTCGATCCGCAAAAACCAACCGCTGTGTAGCGGTGGTTTTT 2460
Db 1262 AAAAAGAGTTCGTAGTCTTCGATCCGCAAAAACCAACCGCTGTGTAGCGGTGGTTTTT 1203
Qy 2461 TGTTCGAAGCAGCAGATTACCGCGCAGAAAAAAGGATCTCAAGAAGATCTCTTTCATCTT 2520
Db 1202 TGTTCGAAGCAGCAGATTACCGCGCAGAAAAAAGGATCTCAAGAAGATCTCTTTCATCTT 1143
Qy 2521 TTCTACGGGTCGTGACGCTCAGTGGAGGAAACTCAGTTTAAAGGATTTTGGTCATGAG 2580
Db 1142 TTCTACGGGTCGTGACGCTCAGTGGAGGAAACTCAGTTTAAAGGATTTTGGTCATGAG 1083
Qy 2581 ATTATCAAAAAAGGATCTTTCACCTAGATCTTTCACCTTAAATTAATAAAGTAAAAATCAAT 2640
Db 1082 ATTATCAAAAAAGGATCTTTCACCTAGATCTTTCACCTTAAATTAATAAAGTAAAAATCAAT 1023
Qy 2641 CTAAGATATATATGAGTAAACTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGAGGAC 2700
Db 1022 CTAAGATATATATGAGTAAACTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGAGGAC 963
Qy 2701 TATCTCAGCGATCTGTCTATTTTCTCATCCATAGTTGCTGATCCCGCTCGTGTAGAT 2760
Db 962 TATCTCAGCGATCTGTCTATTTCTCATTCATAGTTGCTGATCCCGCTCGTGTAGAT 903
Qy 2761 AACTACGATACGGAGGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGCAGAGACC 2820
Db 902 AACTACGATACGGAGGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGCAGAGACC 843
Qy 2821 ACGCTACCGCTCAGATTATCAGCAATAAACAGCAGCGGAGGCGGAGCGGAG 2880
Db 842 ACGCTACCGCTCAGATTATCAGCAATAAACAGCAGCGGAGGCGGAGCGGAG 783
Qy 2881 AAGTGTCTCTCAACTTATCCGCTCCATCCAGTCTATTAATTTGTTCCCGGAAAGCTAG 2940
Db 782 AAGTGTCTCTCAACTTATCCGCTCCATCCAGTCTATTAATTTGTTCCCGGAAAGCTAG 723
Qy 2941 AGTAAGTAGTTTCGCCAGTTAATAGTTTTCGCCAAACGTTGTTGCCATTTGCTACAGGCATCGT 3000
Db 722 AGTAAGTAGTTTCGCCAGTTAATAGTTTTCGCCAAACGTTGTTGCCATTTGCTACAGGCATCGT 663
Qy 3001 GGTGTACGCTCGTGTGTTTGTATGCTTCAATCAGTCCGGTTCCTCCAAAGATCAAGGCG 3060
Db 662 GGTGTACGCTCGTGTGTTTGTATGCTTCAATCAGTCCGGTTCCTCCAAAGATCAAGGCG 603
Qy 3061 AGTTACATGATCCCGCATGTTTGCABAAAAGCGGTTAGTCTTCCGCTCCGATCGT 3120
Db 602 AGTTACATGATCCCGCATGTTTGCABAAAAGCGGTTAGTCTTCCGCTCCGATCGT 543
Qy 3121 TGTCAAGATTAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCAGCTGCATAATTC 3180
Db 542 TGTCAAGATTAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCAGCTGCATAATTC 483
Qy 3181 TCTTACTGTCTATGCGATCCGTAAGATGCTTTTCTGTGATCGGTGAGTACTCAACCAAGTC 3240
Db 482 TCTTACTGTCTATGCGATCCGTAAGATGCTTTTCTGTGATCGGTGAGTACTCAACCAAGTC 423
Qy 3241 ATTCTGAGAAATAGTATGCGCGCAGCCGAGTTGCTCTTCCCGGCTCAATACCGGATAA 3300
Db 422 ATTCTGAGAAATAGTATGCGCGCAGCCGAGTTGCTCTTCCCGGCTCAATACCGGATAA 363


```
QY 3301 TACCGCGCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAGCGTTCTTCGGGGCG 3360
Db |||||||
QY 3362 TACCGCGCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAGCGTTCTTCGGGGCG 303
Db |||||||
QY 3361 AAAACTCTCAAGATCTTACCGCTGTAGATCCAGTTCGATGTAAACCACTCGTGACCC 3420
Db |||||||
QY 302 AAACTCTCAAGATCTTACCGCTGTAGATCCAGTTCGATGTAAACCACTCGTGACCC 243
Db |||||||
QY 3421 CAACTGATCTTCAGCATCTTTTACTTTCCACGAGCTTCTCGGTGAGCAAAAACAGGAAG 3480
Db |||||||
QY 242 CAACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTCTCGGTGAGCAAAAACAGGAAG 183
Db |||||||
QY 3481 GCAAAATCGCCAAAACAGGAATAGCGCGACACGGAATGTTGAATCTCATCTCTT 3540
Db |||||||
QY 182 GCAAAATCGCCAAAACAGGAATAGCGCGACACGGAATGTTGAATCTCATCTCTT 123
Db |||||||
QY 3541 CTTTTTCAATATTATTGAAGCAATTATCAGAGTTATTCTCTCATGAGCGGATACATATT 3600
Db |||||||
QY 122 CTTTTTCAATATTATTGAAGCAATTATCAGAGTTATTCTCTCATGAGCGGATACATATT 63
Db |||||||
QY 3601 TGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGGAAAAAGTGCC 3660
Db |||||||
QY 62 TGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGGAAAAAGTGCC 3
Db |||||||
QY 3661 AC 3662
Db ||
2 AC 1

RESULT 8
US-10-206-030-4/c
; Sequence 4, Application US/10206030
; Publication No. US20020177160A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Lindbo, John A.
; APPLICANT: Fitzmaurice, Wayne P.
; TITLE OF INVENTION: A Method of Increasing Complementarity
; FILE REFERENCE: In A Heteroduplex
; CURRENT APPLICATION NUMBER: US/10/206,030
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: US/10/066,390
; PRIOR FILING DATE: 2002-02-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-206-030-4

Query Match 72.8%; Score 2664.8; DB 13; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

QY 1 CTAATTTGTAAGCGTTAATTTTGTAAATTCGGTTAAATTTTGTAAATCAGCTC 60
Db |||||||
QY 3637 CTAATTTGTAAGCGTTAATTTTGTAAATTCGGTTAAATTTTGTAAATCAGCTC 3578
Db |||||||
QY 61 ATTTTAAACCAATAGCGCGAAATCGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 120
Db |||||||
QY 3577 ATTTTAAACCAATAGCGCGAAATCGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 3518
Db |||||||
QY 121 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGGACTC 180
Db |||||||
QY 3517 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGGACTC 3458
Db |||||||
QY 181 CAACTGCAAGGGCGAAAAACCGTCTATCAGGGCGATGCGCCACTACGTGAACCATCACC 240
Db |||||||
QY 3457 CAACTGCAAGGGCGAAAAACCGTCTATCAGGGCGATGCGCCACTACGTGAACCATCACC 3398
Db |||||||
```

```
QY 241 CTAATCAAGTTTTTTGGGTGCGAGTGCCTGTAAGCACTAAATCGGAACCTTAAAGGAG 300
Db |||||||
QY 3397 CTAATCAAGTTTTTTGGGTGCGAGTGCCTGTAAGCACTAAATCGGAACCTTAAAGGAG 3338
Db |||||||
QY 301 CCCCGCATTTAGAGCTTGACGGGAAAGCCGCGACGTTGGCGAGAAAGGAGGGAAGAA 360
Db |||||||
QY 3337 CCCCGCATTTAGAGCTTGACGGGAAAGCCGCGAAACGTTGGCGAGAAAGGAGGGAAGAA 3278
Db |||||||
QY 361 AGCGAAAGAGAGCGGGCGCTAGGGCGCTGCAAGTGTAGCGGTCACTGCGCCCTAAACCCAC 420
Db |||||||
QY 3277 AGCGAAAGAGAGCGGGCGCTAGGGCGCTGCAAGTGTAGCGGTCACTGCGCGTTAACCCAC 3218
Db |||||||
QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCCATTTCCCATTTCAAGGTGCG 480
Db |||||||
QY 3217 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCCATTTCCCATTTCAAGGTGCG 3158
Db |||||||
QY 481 CAACTGTTGGGAAGGGCGATCGGTGCGGCGCTTTCCGCTATTACGCCAGCTGCGGAAGG 540
Db |||||||
QY 3157 CAACTGTTGGGAAGGGCGATCGGTGCGGCGCTTTCCGCTATTACGCCAGCTGCGGAAGG 3098
Db |||||||
QY 541 GGGATGTGCTCAAGCGCATTAAGTTGGGTAAACGCCAGCGGTTTCCCGAGTCACGAGTTG 600
Db |||||||
QY 3097 GGGATGTGCTCAAGCGCATTAAGTTGGGTAAACGCCAGCGGTTTCCCGAGTCACGAGTTG 3038
Db |||||||
QY 601 TAAAAACGACGGCCAGTGAGCGCGCTCGTTCAATTCACTGTTTTTGAACCCGTTGGAGACGG 660
Db |||||||
QY 3037 TAAAAACGACGGCCAGTGAGCGCGCTTAATACGACTCACTATAGGGCGAAATGGAGCTCCA 2978
Db |||||||
QY 661 GCAGACTCGCGTGCMAATGTGTTTTTACAGCGTGATGGAGCAGATGAAGATGCTCCACAC 720
Db |||||||
QY 2977 CCGCGTGGCGG-----CCGCTCTAGAACTAGTGGATCCCC--CGG 2939
Db |||||||
QY 721 GCTGCAGAACACGCGAGCTAGATTAAACCTPAGAAAGATAATCATATTGTGACGTCAGCTTAA 780
Db |||||||
QY 2938 GCTGCAGGAATTTCTTATTGTATAGTTTCAATGCCATGTGTAAATCCAGCAGAGTTA 2879
Db |||||||
QY 781 AGATAATCATGCGTAAAAATTGACGATGGGATCTGTAAATACGACTCACTATATAGGGCGAAT 840
Db |||||||
QY 2878 CAAACTCAAGAAAGGACCATGTGCTCGCTTTTGGTGGGATCTTTTCGAAAAGGCGAGATT 2819
Db |||||||
QY 841 TGGGTACCGGGCGCCCCCTCGAGGTCGAGGTCGAGGATCGATAAGCTTGTATATCGAATTCCTGC 900
Db |||||||
QY 2818 GTGTCGACAGGTAATGTGTTGTCTGTGTAAAGACAGCGGCCATCGCCAAATGGAGTATTTT 2759
Db |||||||
QY 901 AGCCCGGGGATCCACTAGTTCTAGAGCGCGCGCACCGCGTGGAGCTCCAGCTTTTGT 960
Db |||||||
QY 2758 GTTGATTAATGTTCTGCTAGTTGAAACGATCCATCTTCAATGTTGTGGCGAATTTTGAAGT 2699
Db |||||||
QY 961 TCCCTTTTATGAGGGTTAATTTAGATCCCATGCGTCAATTTTACGCAAGACTATCTTTCTAG 1020
Db |||||||
QY 2698 TAGCTTTGATTCCATTTCTTTTGTGTCGCGTGTATGATATACATTTGTGTAG--TTATAG 2640
Db |||||||
QY 1021 GGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCC 1080
Db |||||||
QY 2639 TTGTACTCGAGTTTGTGTCGCGAATGTTTCCATCTTCTTTAAATCAATACCTTTTAAAC 2580
Db |||||||
QY 1081 CAACTGCGCGTATCTGGGCATCGGGGAGGAAGACCGCTGCTTTTCCCGGAGGTG 1140
Db |||||||
QY 2579 TCGATACGATTAACAGGGTATCACCTTCAAACTTGACTTCAGCAGCGCTCTGTAGTTTC 2520
Db |||||||
QY 1141 AAGCGCGATGGAAGAGTTTGGCGAGGATGACTGCTGCTGCATTGACGTTGTAGCGGAAC 1200
Db |||||||
QY 2519 CGTCACTCTTTGAAAGATATAGTGCCTTCTGTACATAACCTTCGGGCATGGCACTCTTG 2460
Db |||||||
QY 1201 GCACTTTTACCATGATGATTCGGGAAGGTGTGGCCATGCAAGCGCTTTTAAGGTTGAAGTGT 1260
Db |||||||
QY 2459 AAAAAAGTCATGCGCTTTTCAATATGATCCGATTAACGGGAAAGCATTTGAACACCATTAAGAG 2400
Db |||||||
QY 1261 TCGTTTCAGGCCACCTGGGATACAGTTGCTCGCGGCTTTTCCGGACACAGTTTCCCGATG 1320
Db |||||||
QY 2399 AAAGTAGTGACAAAGTGTGGCCCATGGAACAGGTAGTGTTCAGTAGTGCAAAATAAATTTA 2340
Db |||||||
```

QY 1321 TCAGCCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCGGAAGTCCCGTGCCG 1380
Db 2339 AGGGTAAGCTTCCGATGTAGCATCACCTTCCACCTCTCCACTGACAGAAAATTGTGCG 2280
QY 1391 GTGTGCAGATTAAAGACAGCGGTGGCGCTGGGATATTACGTACGACGAGGACGGGTATC 1440
Db 2279 CCAATTAACATCACCATCTAAATTCAACAAGAAATGGGACAACTCCAGTGAAGAAATCTTCT 2220
QY 1441 CTGGCTGATCCGCGAGAAATGGACATGGATACCCCGTGAGTTACCCGCGGGCGCGTT 1500
Db 2219 C---CTTTACTCATCGGTACCCAGCTTTTGTGTCCCTTTAGTGAGGGTTAAATTGCGCGCTT 2163
QY 1501 GCGGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTTCCACA 1560
Db 2162 GCGGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTTCCACA 2103
QY 1561 CAACATACGACCGGAAGCATAAAGTGTAAAGCTTGGGGTGCTTAATGAGTGAGCTAACT 1620
Db 2102 CAACATACGACCGGAAGCATAAAGTGTAAAGCTTGGGGTGCTTAATGAGTGAGCTAACT 2043
QY 1621 CACATTAATTGCGTTTGGCTCACTGCCCGCTTTCAGTTCGGGAACCTGTGTCAGCT 1680
Db 2042 CACATTAATTGCGTTTGGCTCACTGCCCGCTTTCAGTTCGGGAACCTGTGTCAGCT 1983
QY 1681 GCATTAATGAATCGGCCAACGCGGGGAGAGCGGTTTGGCTATTGGGCGCTCTTCCGC 1740
Db 1982 GCATTAATGAATCGGCCAACGCGGGGAGAGCGGTTTGGCTATTGGGCGCTCTTCCGC 1923
QY 1741 TTCTCGCTCACTGACTCGCTCGCTCGGTGCTGCGCTGCGGTCGGGAGCGGTATCAGCTCA 1800
Db 1922 TTCTCGCTCACTGACTCGCTCGCTCGGTGCTGCGGTCGGGAGCGGTATCAGCTCA 1863
QY 1801 CTCAAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATTAACGAGGAAAGAACATGTG 1860
Db 1862 CTCAAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATTAACGAGGAAAGAACATGTG 1803
QY 1861 AGCAAAAGCGCCAGCAAAAGGCCAGGAACCGTAAAGGCCGCTTGTGCGGTTTTTCCA 1920
Db 1802 AGCAAAAGCGCCAGCAAAAGGCCAGGAACCGTAAAGGCCGCTTGTGCGGTTTTTCCA 1743
QY 1921 TAGGCTCGGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1980
Db 1742 TAGGCTCGGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1683
QY 1981 CCCGACAGGACTATAAAGATACAGCGGTTTTCCCTCGAAGCTCCCTCGTGGCTCTCC 2040
Db 1682 CCCGACAGGACTATAAAGATACAGCGGTTTTCCCTCGAAGCTCCCTCGTGGCTCTCC 1623
QY 2041 TGTTCCGACCTGCGCTTACCGGATACCTGTGCGCCTTTCTCCCTTCGGGAAGCGTGCC 2100
Db 1622 TGTTCCGACCTGCGCTTACCGGATACCTGTGCGCCTTTCTCCCTTCGGGAAGCGTGCC 1563
QY 2101 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTGCGTGTAGGTGCTGCTCGAAGCT 2160
Db 1562 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTGCGTGTAGGTGCTGCTCGAAGCT 1503
QY 2161 GGGCTGTGTGACGAACCCCGCTTACGCGGACCGCTGCGCTTATTCGGGTAACTATCG 2220
Db 1502 GGGCTGTGTGACGAACCCCGCTTACGCGGACCGCTGCGCTTATTCGGGTAACTATCG 1443
QY 2221 TCTTGAGTCCAAACCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCACTGGTAAACAG 2280
Db 1442 TCTTGAGTCCAAACCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCACTGGTAAACAG 1383
QY 2281 GATTAGCAGCGAGGTATGTAGGGGTGTACAGAGTTCTTGAAGTGGTGGCTTAAGT 2340
Db 1382 GATTAGCAGCGAGGTATGTAGGGGTGTACAGAGTTCTTGAAGTGGTGGCTTAAGT 1323
QY 2341 CCGCTTACACTAGAAGACAGTATTGTGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 2400
Db 1322 CCGCTTACACTAGAAGACAGTATTGTGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 1263
QY 2401 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCAACCGCTGGTAGCGGTGTTTTT 2460

Db 1262 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCAACCGCTGGTAGCGGTGTTTTT 1203
QY 2461 TGTTTGAAGCAGCAGATTACGCGCAGAAAAGGATCTCAAGNAGATCTCTTTGATCTT 2520
Db 1202 TGTTTGAAGCAGCAGATTACGCGCAGAAAAGGATCTCAAGNAGATCTCTTTGATCTT 1143
QY 2521 TTCTACGGGGTCTGACGCTCAGTGGAAACAAACCTCAGCTTAAAGGATTTTGGTCAATGAG 2580
Db 1142 TTCTACGGGGTCTGACGCTCAGTGGAAACAAACCTCAGCTTAAAGGATTTTGGTCAATGAG 1083
QY 2581 ATTATCAAAAAGGATCTTCACTAGATCTCTTTTAAATTTAAATTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGGATCTTCACTAGATCTCTTTTAAATTTAAATTTAAATCAAT 1023
QY 2641 CTAAAGTATATATGAGTAAACCTTGGTCTGACAGTTACCAATGCTTAAATCAGTGAGGCACC 2700
Db 1022 CTAAAGTATATATGAGTAAACCTTGGTCTGACAGTTACCAATGCTTAAATCAGTGAGGCACC 963
QY 2701 TATCTCAGCGATCTGTCTATTATTCGTTTCAATCCATAGTTGCCTGACTCCCGCTGCTAGAT 2760
Db 962 TATCTCAGCGATCTGTCTATTATTCGTTTCAATCCATAGTTGCCTGACTCCCGCTGCTAGAT 903
QY 2761 AACTACGATACGGAGGGCTTACCATCTGGCCCGCAGTCTGCAATGATACCGCAGAGACC 2820
Db 902 AACTACGATACGGAGGGCTTACCATCTGGCCCGCAGTCTGCAATGATACCGCAGAGACC 843
QY 2821 ACGCTCACGGCTCCAGATTTTATCAGCAATAAACAGCCAGCGGAGCGCGAGCGAG 2880
Db 842 ACGCTCACGGCTCCAGATTTTATCAGCAATAAACAGCCAGCGGAGCGCGAGCGAG 783
QY 2881 AAGTGGTCTCTCAACTTTTATCCGCTCCATCCAGTCTATTAAATTTGTTCGCGGAAAGCTAG 2940
Db 782 AAGTGGTCTCTCAACTTTTATCCGCTCCATCCAGTCTATTAAATTTGTTCGCGGAAAGCTAG 723
QY 2941 AGTAAGTAGTTGCGCAGTTAATAGTTTGGCGAAAGCTTGGTCCCATTTGCTACAGGCATCGT 3000
Db 722 AGTAAGTAGTTGCGCAGTTAATAGTTTGGCGAAAGCTTGGTCCCATTTGCTACAGGCATCGT 663
QY 3001 GGTGTACGCTCGTCTGTTTGTATGGCTTCAATCAGCTCCGTTCCCAACGATCAAGGCG 3060
Db 662 GGTGTACGCTCGTCTGTTTGTATGGCTTCAATCAGCTCCGTTCCCAACGATCAAGGCG 603
QY 3061 AGTTACATGATCCCCCATGTTTGTGCAAAAAGCGGTAGCTCTTCCGTCCTCCGATCGT 3120
Db 602 AGTTACATGATCCCCCATGTTTGTGCAAAAAGCGGTAGCTCTTCCGTCCTCCGATCGT 543
QY 3121 TGTCAAGATTAAGTTGGCCGAGTGTATCACTCATGTTATGGCAGCAGCTGCATAATTC 3180
Db 542 TGTCAAGATTAAGTTGGCCGAGTGTATCACTCATGTTATGGCAGCAGCTGCATAATTC 483
QY 3181 TCTTACTGTATGCTATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCAAGTC 3240
Db 482 TCTTACTGTATGCTATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCAAGTC 423
QY 3241 ATTCTGAGAAATAGTGTATGCGCGACCGAGTTGCTCTTGGCCGCGCTCAATACGGGATAA 3300
Db 422 ATTCTGAGAAATAGTGTATGCGCGACCGAGTTGCTCTTGGCCGCGCTCAATACGGGATAA 363
QY 3301 TACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTTGGAAAAACGTTCTTCGGGCGG 3360
Db 362 TACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTATGAAAAACGTTCTTCGGGCGG 303
QY 3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTGAGTAAACCACTCGTGCAACC 3420
Db 302 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTGAGTAAACCACTCGTGCAACC 243
QY 3421 CAACCTGATCTTTCAGCATCTTTTACTTTTCCACAGCGTTTCTCGGTGAGCAAAAACAGGAAG 3480
Db 242 CAACCTGATCTTTCAGCATCTTTTACTTTTACCAGCGTTTCTCGGTGAGCAAAAACAGGAAG 183
QY 3481 GCAAAATGCCGCAAAAAGGGAATAAGGGCGACAGGAAATGTTGAATACTCATACTCTT 3540

Db 182 GCATAATCGCCGCAAAACGGGAATAGGGCGGACACGGAAATGTTGAATCACTACTCTT 123
Qy 3541 CTTTTCATATTAATGAAGCATTTATCAGGGTTATTTCTCATGAGCGGATACATATT 3600
Db 122 CTTTTCATATTAATGAAGCATTTATCAGGGTTATTTCTCATGAGCGGATACATATT 63
Qy 3601 TGAATGTTATTAAGAAATTAACAAATAGGGTTTCGGCGACATTTCCCGGAAAGTGCC 3660
Db 62 TGAATGTTATTAAGAAATTAACAAATAGGGTTTCGGCGACATTTCCCGGAAAGTGCC 3
Qy 3661 AC 3662
Db 2 AC 1
RESULT 9
US-10-211-079-4/c
; Sequence 4, Application US/10211079
; Publication No. US20030148315A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Vaezhongs, Andrew A.
; APPLICANT: Vojdani, Fakhrieh S.
; APPLICANT: Smith, Mark L.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding CEL I
; FILE OF INVENTION: Endonuclease and Methods of Use Thereof
; FILE REFERENCE: P-LG 5381
; CURRENT APPLICATION NUMBER: US/10/211,079
; PRIOR FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 10/098,155
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-211-079-4
Query Match 72.8%; Score 2664.8; DB 15; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;
Qy 1 CTAATTTGAAGCGTTAATATTTTGTAAATTCGGGTTAAATTTTGTAAATCAGCTC 60
Db 3637 CTAATTTGAAGCGTTAATATTTTGTAAATTCGGGTTAAATTTTGTAAATCAGCTC 3578
Qy 61 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 120
Db 3577 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 3518
Qy 121 GATAGGTTGAGTTGTTCCAGTTTGGAAACAGAGTCACCTATTAAAGAACTGGACTC 180
Db 3517 GATAGGTTGAGTTGTTCCAGTTTGGAAACAGAGTCACCTATTAAAGAACTGGACTC 3458
Qy 181 CAACGTCAAAGGCGGAAACCGTCTATCAGGCGGATGCGCCACTACGTGAACCATCAC 240
Db 3457 CAACGTCAAAGGCGGAAACCGTCTATCAGGCGGATGCGCCACTACGTGAACCATCAC 3398
Qy 241 CTAATCAAGTTTTTGGGTCGAGTGCGGTAAAGCACTAAATTCGGAACCCCTAAAGGGAG 300
Db 3397 CTAATCAAGTTTTTGGGTCGAGTGCGGTAAAGCACTAAATTCGGAACCCCTAAAGGGAG 3338
Qy 301 CCCCAGATTAGAGTTGACGGGGAAAGCCGGGAAACGTTGGCGAGAAAGGAGGAGAA 360
Db 3337 CCCCAGATTAGAGTTGACGGGGAAAGCCGGGAAACGTTGGCGAGAAAGGAGGAGAA 3278
Qy 361 AGCGAAAGGAGCGGCGCTAGGCGCTGCAAGTGTAGCGGTACGCTCGCGGTAAACCA 420
Db 3277 AGCGAAAGGAGCGGCGCTAGGCGCTGCAAGTGTAGCGGTACGCTCGCGGTAAACCA 3218

Qy 421 CACACCCCGCGCTTAATTCGCGCGCTACAGGGCGGCTCCCATTCGCCATTTCAGGCTGCG 480
Db 3217 CACACCCCGCGCTTAATTCGCGCGCTACAGGGCGGCTCCCATTCGCCATTTCAGGCTGCG 3158
Qy 481 CAACTGTTGGGAAGGCGGATCGGTGCGGCCCTCTTCGCTATTACGCGAGCTGCGGAAAGG 540
Db 3157 CAACTGTTGGGAAGGCGGATCGGTGCGGCCCTCTTCGCTATTACGCGAGCTGCGGAAAGG 3098
Qy 541 GGGATGTCGCAAGGCGGATTAAGTTGGGTAAAGTGGGTAAAGTGGGTAAAGTGGGTAAAGTGGGT 600
Db 3097 GGGATGTCGCAAGGCGGATTAAGTTGGGTAAAGTGGGTAAAGTGGGTAAAGTGGGTAAAGTGGGT 3038
Qy 601 TAAACGACGCGGCGAGTCGCGGCCCTCGTTTCATTACGTTTTTGAACCGGTGGAGACCG 660
Db 3037 TAAACGACGCGGCGAGTCGCGGCCCTCGTTTCATTACGTTTTTGAACCGGTGGAGTCCA 2978
Qy 661 GCAGACTCGCGGTGCAAAATGTTTTCAGCGGTGATGGAGCAGATGAAGATGTCGACAC 720
Db 2977 CCGCGGTGCGG-----CGCTCTAGAACTAGTGGATCCCC--CGG 2939
Qy 721 GCTGCAGAACACGCGAGCTAGTTAAACCCCTAGAAAGATAATCATATTGTGACGTAACGTTAA 780
Db 2938 GCTGCAGAAATTCCTTATTTGTATAGTTTCATCCATGCCATGTGTAAATCCCGACGACGATT 2879
Qy 781 AGATAATCATGCGTAAATTTAGCGCATGGATCTGTAAATACGACTCCTATATAGGGCGAAT 840
Db 2878 CAAACTCAAGAAAGGACCATGTGCTCAGCTTTTCTGTTGGGATCTTTTCGAAAGGCGAGATT 2819
Qy 841 TGGGTACCGGGCCCCCCCCCTCGAGGTCGAGGTATCGATAAGCTTGATATCGAATTCCTGC 900
Db 2818 GTGTCACAGGTAATGTTGTCGTGTAAGGACAGGGCCATCGCCAAATTTGGAGTATTTT 2759
Qy 901 AGCCCGGGGGATCCACTAGTTCTTAGAGCGCGCGCCACCGCGGTGGAGCTCCAGCTTTTGT 960
Db 2758 GTTGATAATGCTGCTAGTTGNAACGATCCATCTTCAATGTTGTGCGCAATTTTGAAGT 2699
Qy 961 TCCCTTTAGTAGGGTTAATAGATCCCATGGTCAATTTTACGCGAGACTATCTTTCTAG 1020
Db 2698 TAGCTTTGATTCATCTTTTGTCTGCGGTGATGATACATTTGTGTGAG--TTATAG 2640
Qy 1021 GGTAACTTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTTCGGCTCAGTCATCGCC 1080
Db 2639 TTGTAATCTCAGTTTGTGTCGAGAAATGTTTCCATCTCTTTAAATCAATACCTTTTAAC 2580
Qy 1081 CAACTGCGCGCTATCTGGGCATCGGGGAGGAAGAACCCGCTGCTTTTCCCGGAGGTG 1140
Db 2579 TCGATACGATTAAACAGGGTATCACCTTCAAACTTGACTTCAGCAGCGCTTGTAGTTC 2520
Qy 1141 AAGCGCATGGAAGAGATTTCGCGAGGATGACTGTGCTGCAATTGACGTTGACGGAAC 1200
Db 2519 CCGTCATCTTTGAAAGATATAGTGGGTTCTCTGTACATAACCTTCGGGCATGGCACTTTG 2460
Qy 1201 GCAGTTTTACCATGATGATTCGGGAAGGTGCGCCATGACGCTTTAAACGGTGAACGTG 1260
Db 2459 AAAAAAGTCATGCGTTTCATATGATCCGGATTAACGGGAAAGCAATTTGAACACCATAGAG 2400
Qy 1261 TCGTTTCAGCCCACTCGGATACAGTTTCGTGCGGCTTTTTCGGGACACAGTTTCGGGATG 1320
Db 2399 AAAGTAGTGACAGTTGTTGGCCATGGAACAGGTAGTTTTCAGTAGTGCAAAATTA 2340
Qy 1321 TCAGCCCGAAGCGCATACGAAACCCGAAACAATAACGCGGAGACGCGGAACTGCGGTGCG 1380
Db 2339 AGGGTAAGCTTTCCGTTATGTAGCATCACCTTCAACCTCTCCCATGACGAAATTTGTGC 2280
Qy 1381 GTGTCAGATTATGACAGCGGTGCGGCTGGGATTTACGTCAGGAGGACGGGTATC 1440
Db 2279 CCATTAACATCACCATTCTAAATTCACAAGAAATTTGGGACAACTCCAGTGAAGTTCTTCT 2220
Qy 1441 CTGGCTGGAATCGCGAGAAATGGAATGATGATACCCCGTCAGTTTACCCGCGGCGCGCTT 1500
Db 2219 C---CTTTACTCATCGTACCCAGCTTTTGTTCCTTTTGTAGTGAGGTTTAAATGCGGCTT 2163
Qy 1501 GCGTAAATCATGGTCATAGCTGTTTCTGTGTGAAATTTGTTATTCGCTCACAATTCACAC 1560

Db 2162 |||||GGCGTAATCATGGTCATAGCTGTTTCTCTGTGTGAAATGTTTATCCGCTCACAAATTCACAC 2103
Qy 1561 CAACATACGAGCCGGAAGCAATAAGTGTAAAGCCCTGGGGTGCCTAAATGAGTGAGCTAACT 1620
Db 2102 CAACATACGAGCCGGAAGCAATAAGTGTAAAGCCCTGGGGTGCCTAAATGAGTGAGCTAACT 2043
Qy 1621 CACATTAATTCGGTTGCGCTCACTGCGCGCTTCCAGTTCGGGAACCTGTGCTGCCAGCT 1680
Db 2042 CACATTAATTCGGTTGCGCTCACTGCGCGCTTCCAGTTCGGGAACCTGTGCTGCCAGCT 1983
Qy 1681 GCATTAATGAATTCGGCCAAACCGCGGGGAGAGGGCGTTTGGTGTATGGGGCGCTCTTCGCG 1740
Db 1982 GCATTAATGAATTCGGCCAAACCGCGGGGAGAGGGCGTTTGGTGTATGGGGCGCTCTTCGCG 1923
Qy 1741 TTCCTCGCTCACTGACTCGCTGCGCTCGGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1800
Db 1922 TTCCTCGCTCACTGACTCGCTGCGCTCGGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1863
Qy 1801 CTCAAAGCGGTAATACGGTTTATCCACAGAATCAGGGGATAAACGAGGAAAGAAACATGTG 1860
Db 1862 CTCAAAGCGGTAATACGGTTTATCCACAGAATCAGGGGATAAACGAGGAAAGAAACATGTG 1803
Qy 1861 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAGAGCCGCGTTGCTGCGGTTTTTCCA 1920
Db 1802 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAGAGCCGCGTTGCTGCGGTTTTTCCA 1743
Qy 1921 TAGGCTCGGCCCGCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGTGCGGAA 1980
Db 1742 TAGGCTCGGCCCGCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGTGCGGAA 1683
Qy 1981 CCCGACAGGACTATAAAGATACACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCC 2040
Db 1682 CCCGACAGGACTATAAAGATACACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCC 1623
Qy 2041 TGTTCCGACCCCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTTCGGGAAGCGTGCC 2100
Db 1622 TGTTCCGACCCCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTTCGGGAAGCGTGCC 1563
Qy 2101 GCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAGCT 2160
Db 1562 GCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAGCT 1503
Qy 2161 GGGCTGTGTGACGAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAACTATCG 2220
Db 1502 GGGCTGTGTGACGAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAACTATCG 1443
Qy 2221 TCTTGAGTCCAAACCCGGTAAACACGACTTATCGCACTGGCAGCAGCACTGGTAAACAG 2280
Db 1442 TCTTGAGTCCAAACCCGGTAAACACGACTTATCGCACTGGCAGCAGCACTGGTAAACAG 1383
Qy 2281 GATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAGTGGTGGCTTAACCTA 2340
Db 1382 GATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAGTGGTGGCTTAACCTA 1323
Qy 2341 CGGCTACACTAGAAGACAGTATTGGTATCTGGCTCTGCTGAAGCCAGTATACCTTCGG 2400
Db 1322 CGGCTACACTAGAAGACAGTATTGGTATCTGGCTCTGCTGAAGCCAGTATACCTTCGG 1263
Qy 2401 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGGTTTTTT 2460
Db 1262 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGGTTTTTT 1203
Qy 2461 TGTTTGCAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTTTGATCTTT 2520
Db 1202 TGTTTGCAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTTTGATCTTT 1143
Qy 2521 TTCTACGGGGTCTGACGCTCAGTGGACGAAACTCAGCTTAAAGGATTTTGGTCAATGAG 2580
Db 1142 TTCTACGGGGTCTGACGCTCAGTGGACGAAACTCAGCTTAAAGGATTTTGGTCAATGAG 1083
Qy 2581 ATTATCAAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAAT 2640
|||||

Db 1082 ATTATCAAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAAT 1023
Qy 2641 CTTAAGTATATATCAGTAAACTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGAGGCACC 2700
Db 1022 CTTAAGTATATATCAGTAAACTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGAGGCACC 963
Qy 2701 TATCTCAGCGATCTGTCTATTTCGTTTCATTCATAGTTGCTGACTCCCGTGGTGTAGAT 2760
Db 962 TATCTCAGCGATCTGTCTATTTCGTTTCATTCATAGTTGCTGACTCCCGTGGTGTAGAT 903
Qy 2761 AACTACGATACGGGAGGGCTTACCATCTGGGCCAGTCTGCAATGATACCGGAGACCC 2820
Db 902 AACTACGATACGGGAGGGCTTACCATCTGGGCCAGTCTGCAATGATACCGGAGACCC 843
Qy 2821 ACGCTCACCGCTCCAGATTTTATCAGCAATAAAACAGCAGCCGGAAGGGCCGAGCGAG 2880
Db 842 ACGCTCACCGCTCCAGATTTTATCAGCAATAAAACAGCAGCCGGAAGGGCCGAGCGAG 783
Qy 2881 AAGTGGTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAAATTTGTTCCCGGAAGCTAG 2940
Db 782 AAGTGGTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAAATTTGTTCCCGGAAGCTAG 723
Qy 2941 AGTAAAGTAGTTGCGCAGTTAATAGTTTCCGCAACGTTTGGCATTTGCTACAGGCATCGT 3000
Db 722 AGTAAAGTAGTTGCGCAGTTAATAGTTTTCGCAACGTTTGGCATTTGCTACAGGCATCGT 663
Qy 3001 GGTGTACGCTCGTCTGTTGGTATGGCTTCAATCAGCTCCGGTTCGCCAATCAAGGCG 3060
Db 662 GGTGTACGCTCGTCTGTTGGTATGGCTTCAATCAGCTCCGGTTCGCCAATCAAGGCG 603
Qy 3061 AGTTACATGATPCCCCATGTTGTGCAAAAAAGCGGTAGTCTTCTCGTCTCCGATCGT 3120
Db 602 AGTTACATGATPCCCCATGTTGTGCAAAAAAGCGGTAGTCTTCTCGTCTCCGATCGT 543
Qy 3121 TGTCAAGAGTAGTTGGCCGAGTGTATCACTCATGTTATGCGCAGCACTGCATAATTC 3180
Db 542 TGTCAAGAGTAGTTGGCCGAGTGTATCACTCATGTTATGCGCAGCACTGCATAATTC 483
Qy 3181 TCTTACTGTCTATGCTATGCTTAAGTGTCTTCTGTGATGCTGTGATGCTCAACCAAGTC 3240
Db 482 TCTTACTGTCTATGCTATGCTTAAGTGTCTTCTGTGATGCTGTGATGCTCAACCAAGTC 423
Qy 3241 ATTCTGAGAAATGATGTATGCGCGACCGAGTTGCTCTTTCGCGGCGTCAATACGGGATAA 3300
Db 422 ATTCTGAGAAATGATGTATGCGCGACCGAGTTGCTCTTTCGCGGCGTCAATACGGGATAA 363
Qy 3301 TACCGCGCACATAGCAGAACTTTAAAGTCTCATCTTGGAAACCGTCTTTCGSGGCG 3360
Db 362 TACCGCGCACATAGCAGAACTTTAAAGTCTCATCTTGGAAACCGTCTTTCGSGGCG 303
Qy 3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATTAACCCACTCTCGTGACC 3420
Db 302 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATTAACCCACTCTCGTGACC 243
Qy 3421 CAACTGATCTTTCAGCATCTTTTACTTTTCCACAGCGTTTCTGGGTGAGCAAAACAGGAAG 3480
Db 242 CAACTGATCTTTCAGCATCTTTTACTTTTCCACAGCGTTTCTGGGTGAGCAAAACAGGAAG 183
Qy 3481 GCAAAATGCCCAAAAAAGGGAATAAGGGCAGACGGAATGTTGAATACTCATACTCTT 3540
Db 182 GCAAAATGCCCAAAAAAGGGAATAAGGGCAGACGGAATGTTGAATACTCATACTCTT 123
Qy 3541 CCTTTTCAATATTATTGAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
Db 122 CCTTTTCAATATTATTGAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 63
Qy 3601 TGAATGTATTTAGAAAAATAAAACAAATAGGGGTTCCGCGCACATTTCCCGGAAAGTGCC 3660
Db 62 TGAATGTATTTAGAAAAATAAAACAAATAGGGGTTCCGCGCACATTTCCCGGAAAGTGCC 3
Qy 3661 AC 3662
|||
Db 2 AC 1

RESULT 10

US-10-356-708-17/c

: Sequence 17. Application US/10356708

; Publication No. US20030157682A1

GENERAL INFORMATION:

APPLICANT: Large Scale Biology Corporation

TITLE OF INVENTION: MISMATCH ENDONUCLEASES AND METHODS OF USE THEREOF

; FILE REFERENCE: P-LG 10100

; CURRENT APPLICATION NUMBER: US/10/356,708

; CURRENT FILING DATE: 2003-01-31

; PRIOR APPLICATION NUMBER: 60/353,722

; PRIOR FILING DATE: 2000-02-01

; PRIOR APPLICATION NUMBER: 10/098,155

; PRIOR FILING DATE: 2002-03-14

; PRIOR APPLICATION NUMBER: 10/211,079

; PRIOR FILING DATE: 2002-08-01

; NUMBER OF SEQ ID NOS: 35

```
; SOFTWARE: PatentIn version 3.1
```

; SEQ ID NO 17

; LENGTH: 3637

; TYPE: DNA

; ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Encodes cycle 3 GFP

US-10-356-708-17

Query Match	72.8%;	Score 2664.8;	DB 16;	Length 3637;
Best Local Similarity	83.8%;	Pred. No. 0;		
Matches 3070;	Conservative 0;	Mismatches 567;	Indels 25;	Gaps 4;
Qy	1	CTAAATGTTAAGCGTTAATATTTTGTGTAAATATCGCGTTAAATTTTGTGTAAATCAGCTC	60	
Db	3637	CTAAATGTTAAGCGTTAATATTTTGTGTAAATTCGCGTTAAATTTTGTGTAAATCAGCTC	3578	
Qy	61	ATTTTTTAAACAATAGCGCGAAATCGGCAAAATCCCTTTATAAATCAAAAGATAGACCGA	120	
Db	3577	ATTTTTTAAACAATAGCGCGAAATCGGCAAAATCCCTTTATAAATCAAAAGATAGACCGA	3518	
Qy	121	GATAGGGTTCAGTGTCTGTTCCAGTTTCGGAACAAGAGTCCACTATTAAAGAAACGTGGACTC	180	
Db	3517	GATAGGGTTCAGTGTCTGTTCCAGTTTCGGAACAAGAGTCCACTATTAAAGAAACGTGGACTC	3458	
Qy	181	CAACGTCAAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC	240	
Db	3457	CAACGTCAAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC	3398	
Qy	241	CTAATCAAGTTTTTTTTGGGGTTCGAGTGCCTGTAAAGCACTAAATCGGAACCTTAAGGGAG	300	
Db	3397	CTAATCAAGTTTTTTTTGGGGTTCGAGTGCCTGTAAAGCACTAAATCGGAACCTTAAGGGAG	3338	
Qy	301	CCCCCGATTTAGAGCTTTGACGGGGAAAGCGGGCGAAACGTGGCGAGAAAGGAAGGAAGAA	360	
Db	3337	CCCCCGATTTAGAGCTTTGACGGGGAAAGCGGGCGAAACGTGGCGAGAAAGGAAGGAAGAA	3278	
Qy	361	AGCGAAAGGACGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCAAGTGTGGCGGTACCCAC	420	
Db	3277	AGCGAAAGGACGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCAAGTGTGGCGGTACCCAC	3218	
Qy	421	CACACCCGCGCGCTTAAATGCGCGCTACAGGGCGCGTCCCATTTGCGCAATTCAGGCTGCG	480	
Db	3217	CACACCCGCGCGCTTAAATGCGCGCTACAGGGCGCGTCCCATTTGCGCAATTCAGGCTGCG	3158	
Qy	481	CAACTGTTGGGAAGGGCGATCGGTGCGGGCGCTCTTTCGCTATTACGCCAGCTGCGCAAGG	540	
Db	3157	CAACTGTTGGGAAGGGCGATCGGTGCGGGCGCTCTTTCGCTATTACGCCAGCTGCGCAAGG	3098	
Qy	541	GGGATGTGTCGACGAGCGATTTAGTTGGGTAAAGCCAGGGTTTTCCAGTTCACGACGTTG	600	
Db	3097	GGGATGTGTCGACGAGCGATTTAGTTGGGTAAAGCCAGGGTTTTCCAGTTCACGACGTTG	3038	
Qy	601	TAAAAACGACGGCCAGTGAGGCGCGCTCGTTTCATTTACGTTTTTTTTGAAACCCGTGGAGACGG	660	

Db 1982 GCATTAATGAAATCGGCAACGCGCGGGAGAGGCGGTTTGCGTATTGGCGGCTCTTCGCG 1923
Qy 1741 TTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCA 1800
Db 1922 TTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCA 1863
Qy 1801 CTCAAAGCGGTAATACCGTTATCCACAGAAATCAGGGATACCGAGGAAAGAAATCGTG 1860
Db 1862 CTCAAAGCGGTAATACCGTTATCCACAGAAATCAGGGATACCGAGGAAAGAAATCGTG 1803
Qy 1861 AGCAAAAGCGGAGCAAAAGCGGAGAAACCGTAAAGCGCGGTTGCTGGCGTTTTCOA 1920
Db 1802 AGCAAAAGCGGAGCAAAAGCGGAGAAACCGTAAAGCGCGGTTGCTGGCGTTTTCOA 1743
Qy 1921 TAGGTCGCGCGCGCTGAGGAGCATCAAAATCGACGCTCAAGTCAGAGTGCGGAAA 1980
Db 1742 TAGGTCGCGCGCGCTGAGGAGCATCAAAATCGACGCTCAAGTCAGAGTGCGGAAA 1683
Qy 1981 CCGGACAGGACTATAAAGATACCGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCC 2040
Db 1682 CCGGACAGGACTATAAAGATACCGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCC 1623
Qy 2041 TGTTCCGACCTGCGCTTACCGGATACCGTATCGCTTTCCTTCGGGAAGCGTGCG 2100
Db 1622 TGTTCCGACCTGCGCTTACCGGATACCGTATCGCTTTCCTTCGGGAAGCGTGCG 1563
Qy 2101 GCTTTCTCATAGCTCACGCTGATGATCTCAGTTTCGGTGTAGTCGTTCCGCTCAAGCT 2160
Db 1562 GCTTTCTCATAGCTCACGCTGATGATCTCAGTTTCGGTGTAGTCGTTCCGCTCAAGCT 1503
Qy 2161 GGGCTGTGTGACGAAACCCCGCTTTCAGCGCGGACCGCTCGCGCTTATCCGGTAACTATCG 2220
Db 1502 GGGCTGTGTGACGAAACCCCGCTTTCAGCGCGGACCGCTCGCGCTTATCCGGTAACTATCG 1443
Qy 2221 TCTTGAGTCCAAACCGGTAAAGACACGACTTATTCGCACTGGCAGCAGCACTGGTAAACAG 2280
Db 1442 TCTTGAGTCCAAACCGGTAAAGACACGACTTATTCGCACTGGCAGCAGCACTGGTAAACAG 1383
Qy 2281 GATTAGCAGCGAGGTATGAGCGGTGCTACAGATTCCTGAGTGGTGCGCTTAACTA 2340
Db 1382 GATTAGCAGCGAGGTATGAGCGGTGCTACAGATTCCTGAGTGGTGCGCTTAACTA 1323
Qy 2341 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTCGG 2400
Db 1322 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTCGG 1263
Qy 2401 AAAAGAGTTGGTAGCTCTTGATCCGGCAAAACCAACCGCTGTGTAGCGGTGGTTTTTT 2460
Db 1262 AAAAGAGTTGGTAGCTCTTGATCCGGCAAAACCAACCGCTGTGTAGCGGTGGTTTTTT 1203
Qy 2461 TGTTTGAAGCAGCAGATTACCGCAGAAAGGATCTCAAGAGATCCTTTGATCTT 2520
Db 1202 TGTTTGAAGCAGCAGATTACCGCAGAAAGGATCTCAAGAGATCCTTTGATCTT 1143
Qy 2521 TTCTACCGGGTCTGACGCTCAGTGGAAAGGAACTCAGCTTAAAGGATTTTGGTCAATGAG 2580
Db 1142 TTCTACCGGGTCTGACGCTCAGTGGAAAGGAACTCAGCTTAAAGGATTTTGGTCAATGAG 1083
Qy 2581 ATTATCAAAAAGGATCTTTCACCTAGATCCTTTAAATTAATAATGAAGTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGGATCTTTCACCTAGATCCTTTAAATTAATAATGAAGTTTAAATCAAT 1023
Qy 2641 CTAAGATATATAGTAAACTTGTGCTGACAGTTTACCAATGCTTAACTAGTGAGGCACC 2700
Db 1022 CTAAGATATATAGTAAACTTGTGCTGACAGTTTACCAATGCTTAACTAGTGAGGCACC 963
Qy 2701 TATCTCAGCGATCTGCTATTTGTTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGAT 2760
Db 962 TATCTCAGCGATCTGCTATTTGTTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGAT 903
Qy 2761 AACTACGATAGCGGAGGCTTACCATCTGGCCCGCAGTCTGCAATGATACCGGAGACC 2820
Db 902 AACTACGATAGCGGAGGCTTACCATCTGGCCCGCAGTCTGCAATGATACCGGAGACC 843

Qy 2821 AGCTCACCGCTCCAGATTTATCAGCAATAAAACAGCCAGCCGGAAGCGCGAGCGCAG 2880
Db 842 AGCTCACCGCTCCAGATTTATCAGCAATAAAACAGCCAGCCGGAAGCGCGAGCGCAG 783
Qy 2881 AAGTGGTCTCTCAACTTTATCCGCTCCATCCAGTCTATAATTTGTTCCCGGAAAGCTAG 2940
Db 782 AAGTGGTCTCTCAACTTTATCCGCTCCATCCAGTCTATAATTTGTTCCCGGAAAGCTAG 723
Qy 2941 AGTAAAGTAGTTGCGCAGTTAATAGTTTTCGCAACGTTTTCGCAATTCGTACAGGCATCGT 3000
Db 722 AGTAAAGTAGTTGCGCAGTTAATAGTTTTCGCAACGTTTTCGCAATTCGTACAGGCATCGT 663
Qy 3001 GGTGTACGCTCGCTCGTTTGGTATGGCTTCAATCAGCTCCGTTCCCAACGATCAAGGCG 3060
Db 662 GGTGTACGCTCGCTCGTTTGGTATGGCTTCAATCAGCTCCGTTCCCAACGATCAAGGCG 603
Qy 3061 AGTTACATGATCCCCCATGTTTGTGCAAAAGAGCGTTAGTCTCTTCGTTCCGATCGT 3120
Db 602 AGTTACATGATCCCCCATGTTTGTGCAAAAGAGCGTTAGTCTCTTCGTTCCGATCGT 543
Qy 3121 TGTCAAGATAGTTGCGCGCAGTGTATCACTCATGTTATGCGCAGCACTGCATAATTC 3180
Db 542 TGTCAAGATAGTTGCGCGCAGTGTATCACTCATGTTATGCGCAGCACTGCATAATTC 483
Qy 3181 TCTTACTGTCTATGCTATGCTTAAGATGCTTTTCTGTGACTGCTGAGTACTCAACCAAGTC 3240
Db 482 TCTTACTGTCTATGCTTAAGATGCTTTTCTGTGACTGCTGAGTACTCAACCAAGTC 423
Qy 3241 ATTCTGAGAAATAGTGTATGCGCGCACCGAGTTGCTCTTTGCCCGGCGTCAATACGGGATAA 3300
Db 422 ATTCTGAGAAATAGTGTATGCGCGCACCGAGTTGCTCTTTGCCCGGCGTCAATACGGGATAA 363
Qy 3301 TACCGCGCACATAGCAGAACTTTAAAGTCTCATCATTTGGAAACGTTCTTCGGGGCG 3360
Db 362 TACCGCGCACATAGCAGAACTTTAAAGTCTCATCATTTGGAAACGTTCTTCGGGGCG 303
Qy 3361 AAAACTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 3420
Db 302 AAAACTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 243
Qy 3421 CAACGTATCTTCAGATCTTTTACCTTTTCCAGCGGTTTTCAGCGGTTTTCAGGAGCAAG 3480
Db 242 CAACGTATCTTCAGATCTTTTACCTTTTCCAGCGGTTTTCAGCGGTTTTCAGGAGCAAG 183
Qy 3481 GCAAAATGCCCAAAAGGCAATAAGGCGCACGGAATGTTGAATCACTCATACTCTT 3540
Db 182 GCAAAATGCCCAAAAGGCAATAAGGCGCACGGAATGTTGAATCACTCATACTCTT 123
Qy 3541 CCTTTTTCATATTTATTTGAAGCATTTTATCAGGGTTATTTGTTCTCATGAGCGGATACATATT 3600
Db 122 CCTTTTTCATATTTATTTGAAGCATTTTATCAGGGTTATTTGTTCTCATGAGCGGATACATATT 63
Qy 3601 TGAATGTATTTAGAAAAATAAAATAAGGGGTTTCGCGCACATTTTCCCGGAAAGTGCC 3660
Db 62 TGAATGTATTTAGAAAAATAAAATAAGGGGTTTCGCGCACATTTTCCCGGAAAGTGCC 3
Qy 3661 AC 3662
Db 2 AC 1

RESULT 11
US-10-205-772-4/c
; Sequence 4, Application US/10205772
; Publication No. US20030186261A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Lindbo, John A.
; APPLICANT: Fitzmaurice, Wayne P.
; TITLE OF INVENTION: A Method of Increasing Complementarity
; IN A Heteroduplex
; FILE REFERENCE: P-LG 4878

CURRENT APPLICATION NUMBER: US/10/205,772

CURRENT FILING DATE: 2002-07-25

PRIOR APPLICATION NUMBER: US/10/066,390

PRIOR FILING DATE: 2002-02-01

NUMBER OF SEQ ID NOS: 15

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 4

LENGTH: 3637

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: synthetic construct

OTHER INFO
S-10-205-772-4

Query Match 72.8% Score 2664.8 DB 16: Length 3637:

Query Match	92.0%	Score	2004.0
Best Local Similarity	83.8%	Pred. No.	0

Best local similarity 65.0%, freq: NO: 0,
Matches 3070: Conservative 0: Mismatches 567: Indels 25: Gaps 4:

1 CTAAATTGTAGCGTTAATAATTTGTAAATTCTGTTAATCAGCTC 60

Y I C I A A A T G T A A G C G T A A A T A T T T G T T A A A T C C G T T A A A T T T T G T T A A A T C A G C T C 80
|||||

3637 C TAAAT TGTAAAGCGTTAAATAATTTGTGTTAAATTTTGTGTTAAATCAGCTC 3578

61 A TTTT TTTT A C C A A T A C C C C C A A A T C C C A A A A T C A A A A C A A A T A C C A A C C A 120

[illegible]

3577 ATTTTAAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAGAATAGACCGA 3518

121 GATAGGGTTGAGTGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTC 180

100

181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240

b 3457 CAACGTCAAAGGCGAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

3397 CTAATCAAGTCTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCCTAAAGGGAG 3338

301 CCCCCGATTTAGAGCTTTGACGGGGAAAGCCGGCCGAACGTGGCGAGAAAGGAAGCGGAAGAA 360

[illegible]

U
333 / CCCCAGATTTAGAGGCTTGACCGGGGAHHGCCGGCAGAACGTGGCAGAAAGGGAAGGGAAGA 3278

361 AGCGAAGGACGGCGCGTACGGCGCGTCCCACTGTATGCCGTACGGTGCGGTACCAC A20

221 0

3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCAAGTGGCGGTAACAC 3218

421 CACACCGCGCGCTTAATGCGCGCTACAGGCGGCTCCCATTCGCCATTACAGCTGCG 480

[illegible]

481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAGG 540

b 3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098

3097 GGGATGTGCTGCAAGCGATTAAATTGGTAACGCCAGGGTTTCCAGTCACGACGTTG 3038

QY 1921 TAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1980
Db 1742 TAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1683
QY 1981 CCCGACAGACTATAAAGATACACAGCGGTTTCCCTCTGGAAGCTCCCTCGTGGCTCTCC 2040
Db 1682 CCCGACAGACTATAAAGATACACAGCGGTTTCCCTCTGGAAGCTCCCTCGTGGCTCTCC 1623
QY 2041 TGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGC 2100
Db 1622 TGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGC 1563
QY 2101 GCTTTCTCATAGCTCACGCTGPAAGTATCTCAGTTTCGGGTAGGTGCTTCGCTCAAGCT 2160
Db 1562 GCTTTCTCATAGCTCACGCTGPAAGTATCTCAGTTTCGGGTAGGTGCTTCGCTCAAGCT 1503
QY 2161 GGGCTGTGTGACGAAACCCCGGTTACGCCGACCGCTGCGCTTATCCGGTAACTATCG 2220
Db 1502 GGGCTGTGTGACGAAACCCCGGTTACGCCGACCGCTGCGCTTATCCGGTAACTATCG 1443
QY 2221 TCTTGAGTCCAAACCGGTAAAGACAGACTTATCGGCACCTGGCAGCAGCCACTGGTAAACAG 2280
Db 1442 TCTTGAGTCCAAACCGGTAAAGACAGACTTATCGGCACCTGGCAGCAGCCACTGGTAAACAG 1383
QY 2281 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTA 2340
Db 1382 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTA 1323
QY 2341 CGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACCTTCGG 2400
Db 1322 CGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACCTTCGG 1263
QY 2401 AAAAAGAGTTGTTAGTCTTGTATCCGCAAAACAAACACACCGCTGTAGCGGTGGTTTTT 2460
Db 1262 AAAAAGAGTTGTTAGTCTTGTATCCGCAAAACAAACACACCGCTGTAGCGGTGGTTTTT 1203
QY 2461 TGTTTGCAAGCAGCAGATTAACGCGCAGAAAAAAGGATCTCAAGAAGATCTCTTTGATCTT 2520
Db 1202 TGTTTGCAAGCAGCAGATTAACGCGCAGAAAAAAGGATCTCAAGAAGATCTCTTTGATCTT 1143
QY 2521 TTCTACGGGGTCTGACGCTCAGTGGAAGAAACCTCAGTTTAAGGGATTTTGGTCATGAG 2580
Db 1142 TTCTACGGGGTCTGACGCTCAGTGGAAGAAACCTCAGTTTAAGGGATTTTGGTCATGAG 1083
QY 2581 ATTATCAAAAAGGATCTTCACTAGATCTTCTTAAATTAATAAATGAAGTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGGATCTTCACTAGATCTTCTTAAATTAATAAATGAAGTTTAAATCAAT 1023
QY 2641 CTAAGTATATATAGTAAACTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGAGGCACC 2700
Db 1022 CTAAGTATATATAGTAAACTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGAGGCACC 963
QY 2701 TATCTCAGGATCTGTCTATTTCGTTTCATTCATAGTTGCTGACTCCCGCTCGTGTAGAT 2760
Db 962 TATCTCAGGATCTGTCTATTTCGTTTCATTCATAGTTGCTGACTCCCGCTCGTGTAGAT 903
QY 2761 AACTACGATACGGAGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGCAGAGACC 2820
Db 902 AACTACGATACGGAGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGCAGAGACC 843
QY 2821 ACGCTCACCGGCTCCAGATTTTATCAGCAATAAACCAGCAGCCGGAAGGGCGGAGCGAG 2880
Db 842 ACGCTCACCGGCTCCAGATTTTATCAGCAATAAACCAGCAGCCGGAAGGGCGGAGCGAG 783
QY 2881 AAGTGGTCTGCAACTTTATCGGCTTCCATCAGTCTATTAAATTTGTCGGGAAGCTAG 2940
Db 782 AAGTGGTCTGCAACTTTATCGGCTTCCATCAGTCTATTAAATTTGTCGGGAAGCTAG 723
QY 2941 AGTAAGTAGTTCCGAGTTAATAGTTTTCGCAACGTTTGTGCTTGTACAGGCATCGT 3000
Db 722 AGTAAGTAGTTCCGAGTTAATAGTTTTCGCAACGTTTGTGCTTGTACAGGCATCGT 663
QY 3001 GGTGTACGCTCGTCTGTTTGGTATGGCTTCAATTCAGCTCCGGTTCCTCCCAACGATCAAGCGC 3060

Db 662 GGTGTACGCTCGTCTGTTTGGTATGGCTTCAATTCAGCTCCGGTTCCTCCCAACGATCAAGCGC 603
QY 3061 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGCTCCTTCGGTCTCTCCGATCGT 3120
Db 602 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGCTCCTTCGGTCTCTCCGATCGT 543
QY 3121 TGTCAAGAGTAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCAGCTGCATAATTC 3180
Db 542 TGTCAAGAGTAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCAGCTGCATAATTC 483
QY 3181 TCTTACTGTCATGTCATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCAAGTC 3240
Db 482 TCTTACTGTCATGTCATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCAAGTC 423
QY 3241 ATTCTGAGAATAGTGTATGCGCGACCGAGTTGCTTTCGCCGGCTCAATACGGGATAA 3300
Db 422 ATTCTGAGAATAGTGTATGCGCGACCGAGTTGCTTTCGCCGGCTCAATACGGGATAA 363
QY 3301 TACCGCGCCACATAGCAGAACTTTAAAAAGTGTCTCATCTTGGAAAAACGTTCTTCGGGGCG 3360
Db 362 TACCGCGCCACATAGCAGAACTTTAAAAAGTGTCTCATCTTGGAAAAACGTTCTTCGGGGCG 303
QY 3361 AAACTCTCAAGGATCTTACCGCTGTGTGAGATCCAGTTGATGCCCACTCGTGCACC 3420
Db 302 AAACTCTCAAGGATCTTACCGCTGTGTGAGATCCAGTTGATGCCCACTCGTGCACC 243
QY 3421 CAACCTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGTGAGCAAAAACAGGAAG 3480
Db 242 CAACCTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGTGAGCAAAAACAGGAAG 183
QY 3481 GCAAAATCCGCAAAAAAGGAAATAAGGCGCAGACGGAATAATGTTGAATCTCATACTCTT 3540
Db 182 GCAAAATCCGCAAAAAAGGAAATAAGGCGCAGACGGAATAATGTTGAATCTCATACTCTT 123
QY 3541 CTTTTTCAATATTATTGAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
Db 122 CTTTTTCAATATTATTGAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 63
QY 3601 TGAATGTATTTAGAAAAATAACAAATAGGAGTTTCGCCGACATTTCCCGAAAAAGTGCC 3660
Db 62 TGAATGTATTTAGAAAAATAACAAATAGGAGTTTCGCCGACATTTTCGCCAAAAAGTGCC 3
QY 3661 AC 3662
Db 2 AC 1

RESULT 12
US-10-280-913A-17/c
; Sequence 17, Application US/10280913A
; Publication No. US20040110130A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/280,913A
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

OTHER INFORMATION: Encodes cycle 3 GFP
US-10-280-913A-17

Query Match		72.8%;	Score 2664.8;	DB 18;	Length 3637;
Best Local Similarity		83.8%;	Pred. No. 0;		
Matches 3070;		Conservative 0;	Mismatches 567;	Indels 25;	Gaps 4;
QY	1	CTAAATGTAAGCGTTAAATATTTTGTGTAATAATTCGGGTTAAATTTTGTGTAATAATTCAGCTC	60		
DB	3637	CTAAATGTAAGCGTTAAATATTTTGTGTAATAATTCGGGTTAAATTTTGTGTAATAATTCAGCTC	3578		
QY	61	ATTTTAAACCAATAGCGGCAAAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA	120		
DB	3577	ATTTTAAACCAATAGCGGCAAAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA	3518		
QY	121	GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAAGTGGACTC	180		
DB	3517	GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAAGTGGACTC	3458		
QY	181	CAACGTCAAAAGGCGGAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCAC	240		
DB	3457	CAACGTCAAAAGGCGGAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCAC	3398		
QY	241	CTAAATCAAGTCTTTTGGGTCGAGGTCCGTAAGCACTAAATCGGAACCCCTAAAGGGAG	300		
DB	3397	CTAAATCAAGTCTTTTGGGTCGAGGTCCGTAAGCACTAAATCGGAACCCCTAAAGGGAG	3338		
QY	301	CCCCGATTTAGAGCTTTGACGGGAAAGCCGGCAACGTGGCGAGAAAGAAAGGGAAGAA	360		
DB	3337	CCCCGATTTAGAGCTTTGACGGGAAAGCCGGCAACGTGGCGAGAAAGAAAGGGAAGAA	3278		
QY	361	AGCGAAGAGCGGCGCTAGGCGCTGCGCAAGTGTAGCGGTACGCTCGCGTAAACAC	420		
DB	3277	AGCGAAGAGCGGCGCTAGGCGCTGCGCAAGTGTAGCGGTACGCTCGCGTAAACAC	3218		
QY	421	CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCCATTCGCCATTACGGTCCG	480		
DB	3217	CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCCATTCGCCATTACGGTCCG	3158		
QY	481	CAACTGTTGCGAAGGCGATCGGTGGGCGCTCTTCGCTATTACGCCAGCTGGCGAAAGG	540		
DB	3157	CAACTGTTGCGAAGGCGATCGGTGGGCGCTCTTCGCTATTACGCCAGCTGGCGAAAGG	3098		
QY	541	GGGATGTGTCGAAGCGAATTAAGTTGGTAAACGCCAGGCTTTCCAGTACGAGCTTG	600		
DB	3097	GGGATGTGTCGAAGCGAATTAAGTTGGTAAACGCCAGGCTTTCCAGTACGAGCTTG	3038		
QY	601	TAAACGACGCGCAGTACGCGGCTCGTTTCAATCAGTTTGTGAACCCGCTGGAGACGG	660		
DB	3037	TAAACGACGCGCAGTACGCGGCTCGTTTCAATCAGTTTGTGAACCCGCTGGAGACGG	2978		
QY	661	GCAGACTCGCGTGCAAAATGTGTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC	720		
DB	2977	CGCGGTGCGG-----CGGCTCTAGAAGTAGTGATCCCC--CGG	2939		
QY	721	GCTGCAGAACACGCACTAGATTAAACCTAGAAAGATAATCATATTGTGACGTACGTAA	780		
DB	2938	GCTGCAGGAATCTTTATTGTATGATTTCATCCATGCCATGTGTAATCCCGAGCAGATT	2879		
QY	781	AGATAATCATGCGTAAATATGAGCGATGGATCTGTAAATACGACTCACTATAGGGCGAAT	840		
DB	2878	CAAATCTCAAGAAAGGACCATGTGCTACGCTTTTCGTTGGGATCTTTTCGAAAGGGCAGATT	2819		
QY	841	TGGGTACCGGGCCCCCTCGAGTTCGAGGTATCGATTAAGCTTGATATCGAATTCCTGC	900		
DB	2818	GTGTCACAGGTAATGGTGTCTGGTAAAGGACAGGGCCATCGCCCAATTGGAGTATTTT	2759		
QY	901	AGCCCGGGGATCCACTAGTTCTAGAGCGCGCCGACACCGCGGTGGAGCTCCAGCTTTTGT	960		
DB	2758	GTTGATAATGGTCTGCTAGTTGAAACGATCCATCTTCAATGTTGTGGCGAATTTGAAGT	2699		
QY	961	TCCCTTTAGTAGGGGTAAATTAGATCCCATGCGTCAATTTTACGACAGACTATCTTTCTAG	1020		

DB	2698	TAGCTTTGATCCATTTCTTTGTTGTCGCGGTATGTATATCATTTGTGTGAG-TTATAG	2640
QY	1021	GGTTAATCTAGCTGCATCAGGATCATATCGTCCGGTCTTTTTCGGCTCAGTCATCGCC	1080
DB	2639	TTGTAATCTAGCTGTCGCGAATGTTTCCATCTTCTTTAAATCAATACCTTTTAAAC	2580
QY	1081	CAAGCTGGCGCTATCTGGCATCGGGAGGAGAGCCGCTGCTTTTCCCGGAGGTG	1140
DB	2579	TCGATACGATTAACAAAGGTATCACTTCANAACTTGACTTCAGCAGCGCTTTGTAGTTC	2520
QY	1141	AAGCGCATGGAAGAGTTTGCAGGATGACTGCTGCTGCATTTGACGTTGAGCGAAAC	1200
DB	2519	CCGTATCTTTGAAAGATATAGTGGTCTCTGTATACATACCTTCGGGCATGGCACTTTC	2460
QY	1201	GCACGTTTACATGATGATTCGGGAAAGTGTGGCCATGACGCGTTTAAACGTTAACTGT	1260
DB	2459	AAAAAGTCATGCGCTTTCATATGATCCGGATAACCGGAAAGCAATGAAACCATTAAGAG	2400
QY	1261	TCGTTCAGGCACTGGGATACAGTTCGTCGGGCTTTTCCGGACACAGTTCGGGATGG	1320
DB	2399	AAAGTAGTGACAAGTGTGGCCATGGAAACAGGTAGTTTTCAGTAGTGCAAAATAAATTA	2340
QY	1321	TCAGCCCGAAGCGCATCAGCAACCCGAAACAATAACCGGCGACAGCCGGAACCTGCCGTGCCG	1380
DB	2339	AGGTAAGCTTTCCGTATGTAGCATCACCTTCACCTCTCCACTGACAGAAATTTGTGC	2280
QY	1381	GTGTCAGATTAATGACAGCGGTGGCGCTGGGATATTAACGTACGCGAGGACGGGTATC	1440
DB	2279	CCATTAAACATCACCATCTAATTTCAACAAGATTGGGACAACCTCCAGTGAAGTTCTTCT	2220
QY	1441	CTGGCTGGATGCGGCAAAATGGAATGGAATACCCGTCAGTTTACCCGGCGGCGGCTT	1500
DB	2219	C---CTTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAGGTTAATTTGCGGCTT	2163
QY	1501	GGCGTAATCATGTCATAGCTGTTCTGTGTGAATTTGTTATCCGCTCACAATTCACCA	1560
DB	2162	GGCGTAATCATGTCATAGCTGTTCTGTGTGAATTTGTTATCCGCTCACAATTCACCA	2103
QY	1561	CAACATACGAGCGGAAAGCATAAAGTGTAAAGCTTGGGTCCTAAATGATGAGCTAACT	1620
DB	2102	CAACATACGAGCGGAAAGCATAAAGTGTAAAGCTTGGGTCCTAAATGATGAGCTAACT	2043
QY	1621	CACATTAATGCTGTCGCTCATCTGCCGCTTTCAGTTCGGGAAACCTGTCGTCGAGCT	1680
DB	2042	CACATTAATGCTGTCGCTCATCTGCCGCTTTCAGTTCGGGAAACCTGTCGTCGAGCT	1983
QY	1681	GCATTAATCAATCGCCCAACGCGGGAGAGCGGTTTTCGTTATTTGGGCGCTTTCGCG	1740
DB	1982	GCATTAATCAATCGCCCAACGCGGGAGAGCGGTTTTCGTTATTTGGGCGCTTTCGCG	1923
QY	1741	TTCTCTCGCTCACTGACTCGCTCGCTCGCTTCGGCTTCGGCTCGCGGAGCGGTATCAGTCA	1800
DB	1922	TTCTCTCGCTCACTGACTCGCTCGCTCGCTTCGGCTTCGGCTCGCGGAGCGGTATCAGTCA	1863
QY	1801	CTCAAGGCGGTAAATACGTTATCCAGAAATCAGGGGATTAACGAGGAAAGAAACATGTG	1860
DB	1862	CTCAAGGCGGTAAATACGTTATCCAGAAATCAGGGGATTAACGAGGAAAGAAACATGTG	1803
QY	1861	AGCAAAAGGCGGCAAAAGGCGGCAAGCGTAAAGGCGGCTTCGCTGCGCTTTTTC	1920
DB	1802	AGCAAAAGGCGGCAAAAGGCGGCAAGCGTAAAGGCGGCTTCGCTGCGCTTTTTC	1743
QY	1921	TAGGCTCCCGCCCTTCAGCAGCATCAAAAATTCAGCCTCAAGTCAAGAGTGGCGAA	1980
DB	1742	TAGGCTCCCGCCCTTCAGCAGCATCAAAAATTCAGCCTCAAGTCAAGAGTGGCGAA	1683
QY	1981	CCCGACAGGACTATAAAGATACAGCGTTTCCCTCTGGAAGCTCCTCGTGGCTCTCC	2040
DB	1682	CCCGACAGGACTATAAAGATACAGCGTTTCCCTCTGGAAGCTCCTCGTGGCTCTCC	1623
QY	2041	TGTTCCGACCTTCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAGCGTGC	2100
DB	1622	TGTTCCGACCTTCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAGCGTGC	1563

```
QY 2101 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGTTGAGTTCGTTCCGCTCAAGCT 2160
Db |||||||
QY 1562 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGTTGAGTTCGTTCCGCTCAAGCT 1503
Db |||||||
QY 2161 GGGCTGTGTGCAAGAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGGTAACTATCG 2220
Db |||||||
QY 1502 GGGCTGTGTGCAAGAACCCCGCTTCAGCCCGACCGCTTATCCGGTAACTATCG 1443
QY 2221 TCTTCAGTCCAAACCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAG 2280
Db |||||||
QY 1442 TCTTCAGTCCAAACCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAG 1383
QY 2281 GATTAGCAGAGCGAGGTATGTAGGCGGTCTACAGAGTCTTGAAGTGTGGCCCTAACTA 2340
Db |||||||
QY 1382 GATTAGCAGAGCGAGGTATGTAGGCGGTCTACAGAGTCTTGAAGTGTGGCCCTAACTA 1323
QY 2341 CGGCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTACCTTCGG 2400
Db |||||||
QY 1322 CGGCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTACCTTCGG 1263
QY 2401 AAAAGAGTGTGTAGCTCTTGATCCGGCAAAACCAACCCGCTGCTAGCGGTGGTTTTT 2460
Db |||||||
QY 1262 AAAAGAGTGTGTAGCTCTTGATCCGGCAAAACCAACCCGCTGCTAGCGGTGGTTTTT 1203
QY 2461 TGTTCGAAAGCAGAGATTACCGCAGAGAAAGAGTCTCAAGAGATCTCTTTGATCTT 2520
Db |||||||
QY 1202 TGTTCGAAAGCAGAGATTACCGCAGAGAAAGAGTCTCAAGAGATCTCTTTGATCTT 1143
QY 2521 TTCTACGGGGTCTGACGCTCAGTGTGAACGAAACTCAGCTTAAAGGATTTTGGTTCATGAG 2580
Db |||||||
QY 1142 TTCTACGGGGTCTGACGCTCAGTGTGAACGAAACTCAGCTTAAAGGATTTTGGTTCATGAG 1083
QY 2581 ATTATCAAAAGGATCTTCACTAGATCTTTTAAATTTAAATCAAGTTTAAATCAAT 2640
Db |||||||
QY 1082 ATTATCAAAAGGATCTTCACTAGATCTTTTAAATTTAAATCAAGTTTAAATCAAT 1023
QY 2641 CTAAAGTATATAGTAAACTTGTCTGACAGTTACCAATGCTTTAACTCAGTGAGGCACC 2700
Db |||||||
QY 1022 CTAAAGTATATAGTAAACTTGTCTGACAGTTACCAATGCTTTAACTCAGTGAGGCACC 963
QY 2701 TATCTCAGGGATCTGTCTATTTTCGTTTCATPCATAGTTGCTGACTCCCGCGTGTAGAT 2760
Db |||||||
QY 962 TATCTCAGGGATCTGTCTATTTTCGTTTCATPCATAGTTGCTGACTCCCGCGTGTAGAT 903
QY 2761 AACTACGATACGGGAGGCTTACCATCTGCCCCAGTCTGCAATGATACCGCAGAGACC 2820
Db |||||||
QY 902 AACTACGATACGGGAGGCTTACCATCTGCCCCAGTCTGCAATGATACCGCAGAGACC 843
QY 2821 ACGCTCACCGGCTCCAGATTTTATCAGCAATAAACCCAGCAGCCGGAAGGCGCGAGCGCAG 2880
Db |||||||
QY 842 ACGCTCACCGGCTCCAGATTTTATCAGCAATAAACCCAGCAGCCGGAAGGCGCGAGCGCAG 783
QY 2881 AAGTGGTCTCGAACTTTTATCCGCTCCATCCAGTCTATTAAATGTTCCCGGGAAGCTAG 2940
Db |||||||
QY 782 AAGTGGTCTCGAACTTTTATCCGCTCCATCCAGTCTATTAAATGTTCCCGGGAAGCTAG 723
QY 2941 AGTAAGTATGTCGCCAGTTAATAGTTTTCGGCAACGTTGTTGCCATTGCTACAGGCATCGT 3000
Db |||||||
QY 722 AGTAAGTATGTCGCCAGTTAATAGTTTTCGGCAACGTTGTTGCCATTGCTACAGGCATCGT 663
QY 3001 GGTGTACGCTCGTCTGTTGGTATGGCTTCATTTCAGCTCCGTTCCCAACCATCAAGGCG 3060
Db |||||||
QY 662 GGTGTACGCTCGTCTGTTGGTATGGCTTCATTTCAGCTCCGTTCCCAACCATCAAGGCG 603
QY 3061 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGTCTCTTCGTTCCGATCGT 3120
Db |||||||
QY 602 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGTCTCTTCGTTCCGATCGT 543
QY 3121 TGTCAAGAGTAAGTGGCCGAGTGTATCACTCATGTTTATGGCAGCAGCTGCAATAATTC 3180
Db |||||||
QY 542 TGTCAAGAGTAAGTGGCCGAGTGTATCACTCATGTTTATGGCAGCAGCTGCAATAATTC 483
```

```
QY 3181 TCTTACTGTCTATGTCATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCAAGTC 3240
Db |||||||
QY 482 TCTTACTGTCTATGTCATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCAAGTC 423
QY 3241 ATTCTGAGAAATAGTGTATGCGCGCAGCCGAGTTGCTCTTTGCCCGCGGTCAATACGGGATAA 3300
Db |||||||
QY 422 ATTCTGAGAAATAGTGTATGCGCGCAGCCGAGTTGCTCTTTGCCCGCGGTCAATACGGGATAA 363
QY 3301 TACCGCGCACATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAAAACGTTCTTCGGGGCG 3360
Db |||||||
QY 362 TACCGCGCACATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAAAACGTTCTTCGGGGCG 303
QY 3361 AAAAATCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCACTCGTGCACC 3420
Db |||||||
QY 302 AAAAATCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCACTCGTGCACC 243
QY 3421 CAACTGATCTTCAGCATCTTTTACCTTTCCAGAGCTTTCTGGGTGAGCAAAAACAGGAAG 3480
Db |||||||
QY 242 CAACTGATCTTCAGCATCTTTTACCTTTCCAGAGCTTTCTGGGTGAGCAAAAACAGGAAG 183
QY 3481 GCAAAATCCGCAAAAAGGGAATAAGGGCGACACGGAATAAGTTCGAATACTCATACTCTT 3540
Db |||||||
QY 182 GAAAATCCGCAAAAAGGGAATAAGGGCGACACGGAATAAGTTCGAATACTCATACTCTT 123
QY 3541 CCTTTTCAATATTTATTTGAAGCATTTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
Db |||||||
QY 122 CCTTTTCAATATTTATTTGAAGCATTTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 63
QY 3601 TGAATGTTATTTAGAAAAATAAAATAGGGGTTCCGCGCACATTTCCCGGAAAAAGTGCC 3660
Db |||||||
QY 62 TGAATGTTATTTAGAAAAATAAAATAGGGGTTCCGCGCACATTTCCCGGAAAAAGTGCC 3
QY 3661 AC 3662
Db ||
QY 2 AC 1
```

RESULT 13
US-10-684-134-17/c
; Sequence 17, Application US/10684134
; Publication No. US20040142433A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/684,134
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Encodes cycle 3 GFP
US-10-684-134-17

Query Match 72.8%; Score 2664.8; DB 18; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;
QY 1 CTAATTTCTAGCGTTAATATTTTCTTAAATTCGCGTTAAATTTTCTTAAATCAGCTC 60
Db |||||||
QY 3637 CTAATTTCTAGCGTTAATATTTTCTTAAATTCGCGTTAAATTTTCTTAAATCAGCTC 3578
Db |||||||

QY 61 ATTTTAAACCAATAGGCGCAATCGGCAAAATCCCTTATTAATCAAAAGATAGACCGA 120
Db 3577 ATTTTAAACCAATAGGCGCAATCGGCAAAATCCCTTATTAATCAAAAGATAGACCGA 3518
QY 121 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACTGTGACTC 180
Db 3517 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACTGTGACTC 3458
QY 181 CAAAGTCAAAGGCGGCAAAACCGTCTATCAGGCGATGCGCCCACTACGTGAACCAATCACC 240
Db 3457 CAAAGTCAAAGGCGGCAAAACCGTCTATCAGGCGATGCGCCCACTACGTGAACCAATCACC 3398
QY 241 CTAATCAAGTTTGGGTGAGTGTGCGGTAAAGCACTAAATCGGAAACCTTAAGGGAG 300
Db 3397 CTAATCAAGTTTGGGTGAGTGTGCGGTAAAGCACTAAATCGGAAACCTTAAGGGAG 3338
QY 301 CCCCAGTTTATAGAGTTGACGGGGAAGCGGCAAGCTGGGAGAGAAAGGAAGAA 360
Db 3337 CCCCAGTTTATAGAGTTGACGGGGAAGCGGCAAGCTGGGAGAGAAAGGAAGAA 3278
QY 361 AGCGAAAGAGCGGCGCTAGGCGCTGCGCAAGTGTAGCGGTCAAGCTGCGGTAAACCA 420
Db 3277 AGCGAAAGAGCGGCGCTAGGCGCTGCGCAAGTGTAGCGGTCAAGCTGCGGTAAACCA 3218
QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGCGGCTCCATTGCGCAATTACGCTGCG 480
Db 3217 CACACCGCGCGCTTAATGCGCGCTACAGGCGGCTCCATTGCGCAATTACGCTGCG 3158
QY 481 CAACTGTTGCGAAGGCGAGTCTGCGGCGCTCTTCGCTATTACGCGAGCTGCGGAAAG 540
Db 3157 CAACTGTTGCGAAGGCGAGTCTGCGGCGCTCTTCGCTATTACGCGAGCTGCGGAAAG 3098
QY 541 GGGATGTGCTCAAGCGAGTAAAGTTGGTAAAGCGCAGGCTTTCCAGTACAGAGTTG 600
Db 3097 GGGATGTGCTCAAGCGAGTAAAGTTGGTAAAGCGCAGGCTTTCCAGTACAGAGTTG 3038
QY 601 TAAACGAGCGGAGTACGCGGCTGCTGTTCAATCAGTTTGTGAACCGTGGAGAGCG 660
Db 3037 TAAACGAGCGGAGTACGCGGCTGCTGTTCAATCAGTTTGTGAACCGTGGAGAGCG 2978
QY 661 GCAGACTCGCGTGCAAAATGTGTTTACAGCGTGTAGGAGCAGATGAAGATCTCGACAC 720
Db 2977 CCGCGTGGCGG-----CGGCTAGAACTAGTGGATCCCC--CGG 2939
QY 721 GCTGCAGAACACGCGACTAGATTAAACCTTAGAAAGATAATCATATTGTGACGTACGTTAA 780
Db 2938 GCTGCAGGAATCTTATTGTTGATGTTTCATCCATGCCATGTGTAATCCAGCAGAGTTA 2879
QY 781 AGATAATCATGCTGTAATAATGACGATGGGATCTGTAAATACGACTCACTATAGGGCGAAT 840
Db 2878 CAAACTCAAGAAAGGACCATGTGTCAAGCTTTTCGTTGGGATCTTTGCAAAAGGCGAGATT 2819
QY 841 TGGGTACCGGCGCCCCCTCGAGTTCGAGGTATCGATAAGCTTGATATCGAATTCCTGC 900
Db 2818 GTGTCGACAGGTAATGTTGTTCTGTGTAAAGGACAGGGCCATCGGCCAATTTGAGATATTT 2759
QY 901 AGCCCGGGGGATCCACTAGTTCTTAGAGCGGCGGCCACCGCGGTGGAGCTCCAGCTTTTGT 960
Db 2758 GTTGTAAATGTTCTGCTAGTTGAAACGATCCATCTTCAATGTTGTGGCGAATTTTGAAGT 2699
QY 961 TCCCTTTAGTAGGGTTAAATAGATCCCATGCGTCAATTTTACGCGAGACTATCTTTCTAG 1020
Db 2698 TAGCTTTGATTCATCTTTTGTGTTGCTGCGGTGATGTATACATTTGTGTGAG-TTATAG 2640
QY 1021 GGTAACTAGTGCATCAGGATCATATCGTGGGTCTTTTTCGGCTCAGTCACTCGCC 1080
Db 2639 TTGTACTCGAGTTTGTGTCGAGAAATGTTTTCATCTTCTTAAATCAATCCTTTTAAAC 2580
QY 1081 CAAAGTGGCGCTATCTGGGCGATCGGGGAGGAGAACCGCGCTTTTCCCGCGAGGTTG 1140
Db 2579 TCGATACGATTAAACAGGGTATCACCTTCAAACTTGACTTTCAGCAGCGCTTTGTAGTTC 2520
QY 1141 AAGCGGCATGGAAAGAGTTTGGCGAGGATGACTGCTGCTGCATTGACATTGAGCGGAAAC 1200

Db 2519 CCGTCATCTTTGAAAGATATAGTGCTTCTCTGTACATAAACCTTCGGGCATAGCACTCTTG 2460
QY 1201 GCACGCTTTACCATGATGATTCGGGAAGGTGTGGCCATGACGCGCTTTAAACGGTGAACGTG 1260
Db 2459 AAAAGTCAATCGCTTTTCATATGATCCGGATAACGGGAAAGCAATTGAACACCATTAAGAG 2400
QY 1261 TCGTTAGGCCACCTGGGATACAGTTCGTGCGGCTTTTCCGGACACAGTTCGGGATGG 1320
Db 2399 AAAAGTAGTACAAAGTGTGGCCATGGAAACAGGTAGTTTTTCCAGTAGTGCAGAAATAAATTA 2340
QY 1321 TCAGCCCGAAGCGCATACGAAACCCGAAACAATACCGGCGACAGCGCGAACTGCGGCGCG 1380
Db 2339 AGGTAAGCTTTCCGTATGTAGCATCACCTTCACCTCTCTCACTGACAGAAATTTGTGC 2280
QY 1381 GTGTGCAGATTAAATGACAGCGGTGGCGCTGGGATATTACGTCAGCGAGGACGGGTATC 1440
Db 2279 CCATTAAACATCACCATCTAATTTCAACAAGATTGGGACAACCTCCAGTGAAGTTCTTCT 2220
QY 1441 CTGGCTGGATGCGGAGAAATGGACATGATGATACCCGCTGAGTTACCCGGCGGCGCGCTT 1500
Db 2219 C---CTTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTAATTGCGGCGCT 2163
QY 1501 GCGGTAATCATGCTCATAGCTGTTTCTGTGTGAATTCGTATCCGCTCACAAATCCACA 1560
Db 2162 GCGGTAATCATGCTCATAGCTGTTTCTGTGTGAATTCGTATCCGCTCACAAATCCACA 2103
QY 1561 CAACTATACAGCGCGGAAAGCATAAAGTGTAAAGCTGGGGTGCTTAATGATGAGCTAACT 1620
Db 2102 CAACTATACAGCGCGGAAAGCATAAAGTGTAAAGCTGGGGTGCTTAATGATGAGCTAACT 2043
QY 1621 CACATTAATGCTGTGCGTCACTGCGCTTTCAAGTCGGGAAACCTGTGTCGAGCT 1680
Db 2042 CACATTAATGCTGTGCGTCACTGCGCTTTCAAGTCGGGAAACCTGTGTCGAGCT 1983
QY 1681 GCATTAATGAATCGGCGAACGCGCGGAGAGCGGTTTGCCTGATTTGGGCGCTCTCCGC 1740
Db 1982 GCATTAATGAATCGGCGAACGCGCGGAGAGCGGTTTGCCTGATTTGGGCGCTCTCCGC 1923
QY 1741 TTCTCTCGCTCACTGACTCGCTCGCTCGCTTTCGGCTGCGCGAGCGGTATCAGCTCA 1800
Db 1922 TTCTCTCGCTCACTGACTCGCTCGCTCGCTTTCGGCTGCGCGAGCGGTATCAGCTCA 1863
QY 1801 CTCAAAGCGGTAAATACGCTTATCAAGATCAAGGATTAACGCGGAAAGAAATATGTG 1860
Db 1862 CTCAAAGCGGTAAATACGCTTATCAAGATCAAGGATTAACGCGGAAAGAAATATGTG 1803
QY 1861 AGCAAAAGCGCAGCAAAAGGCGCAGGAAACGTTAAAGGCGCGGTTGCTGCGGTTTTCCA 1920
Db 1802 AGCAAAAGCGCAGCAAAAGGCGCAGGAAACGTTAAAGGCGCGGTTGCTGCGGTTTTCCA 1743
QY 1921 TAGGCTCGCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1980
Db 1742 TAGGCTCGCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1683
QY 1981 CCCGACAGGACTATAAAGATACAGCGGTTTCCCTCTGGAAGCTCCCTCGTGGCTCTCC 2040
Db 1682 CCCGACAGGACTATAAAGATACAGCGGTTTCCCTCTGGAAGCTCCCTCGTGGCTCTCC 1623
QY 2041 TGTTCCGACCTTCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAGCGTGC 2100
Db 1622 TGTTCCGACCTTCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAGCGTGC 1563
QY 2101 GCTTTCTCATAGCTCACTGCTGATGATCTCAGTTTCGTTGATAGGTGCTTTCGCTCCAAGCT 2160
Db 1562 GCTTTCTCATAGCTCACTGCTGATGATCTCAGTTTCGTTGATAGGTGCTTTCGCTCCAAGCT 1503
QY 2161 GGGCTGTGTGACGAAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAACTATCG 2220
Db 1502 GGGCTGTGTGACGAAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAACTATCG 1443
QY 2221 TCTTGAGTCCAAACCGGTTAAGACAGACTTATCGCCACTGCGCAGCAGCCACTGTTAACAG 2280

Db 1442 TCTTGAGTCCAAACCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCACTGGTAAACAG 1383
QY 2281 GATTAGCAGAGAGGTATGTAGGCGGTCTACAGATTCTTGAAGTGGTGGCCCTAACTA 2340
Db 1382 GATTAGCAGAGAGGTATGTAGGCGGTCTACAGATTCTTGAAGTGGTGGCCCTAACTA 1323
QY 2341 CGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTCGG 2400
Db 1322 CGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTCGG 1263
QY 2401 AAAAGAGTTGGTACTCTTGATCCGGCAAAACAAACACCGCTGGTAGGGTGGTTTTTT 2460
Db 1262 AAAAGAGTTGGTACTCTTGATCCGGCAAAACAAACACCGCTGGTAGGGTGGTTTTTT 1203
QY 2461 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAGGATCTCAAGAGATCCTTTGATCTT 2520
Db 1202 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAGGATCTCAAGAGATCCTTTGATCTT 1143
QY 2521 TTCTACGGGGTCTGACGCTCAGTGGAAAGAAACTCAGGTTTAAGGGATTTTGGTCAATGAG 2580
Db 1142 TTCTACGGGGTCTGACGCTCAGTGGAAAGAAACTCAGGTTTAAGGGATTTTGGTCAATGAG 1083
QY 2581 ATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAAGTGGTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAAGTGGTTTAAATCAAT 1023
QY 2641 CTAAAGTATATATGATGATAAATCTGTGCTGACAGTTACCAATGCTTAATCAGTGGGCACC 2700
Db 1022 CTAAAGTATATATGATGATAAATCTGTGCTGACAGTTACCAATGCTTAATCAGTGGGCACC 963
QY 2701 TATCTCAGCGATCTGCTATTTTGGTTCATCCATAGTTTGCCTGACTCCCGCGTGTAGAT 2760
Db 962 TATCTCAGCGATCTGCTATTTTGGTTCATCCATAGTTTGCCTGACTCCCGCGTGTAGAT 903
QY 2761 AACTACGATACGGGAGGGTTACCATCTGGCCCAAGTGTGCAATGATACCGCAGAGACC 2820
Db 902 AACTACGATACGGGAGGGTTACCATCTGGCCCAAGTGTGCAATGATACCGCAGAGACC 843
QY 2821 ACGCTCACCGGCTCAGATTTATCAGCAATAAACAGCAGCCGGAAGCGCGCAGCAG 2880
Db 842 ACGCTCACCGGCTCAGATTTATCAGCAATAAACAGCAGCCGGAAGCGCGCAGCAG 783
QY 2881 AAGTGGTCTCGCACTTTTACCGCTCCATCCAGTCTATTAATTTGTCGGGAAGCTAG 2940
Db 782 AAGTGGTCTCGCACTTTTACCGCTCCATCCAGTCTATTAATTTGTCGGGAAGCTAG 723
QY 2941 AGTAAGTAGTTGCCAGTATATAGTTTGGCMAAGTTGTTGTCATTTGCTACAGGCATCGT 3000
Db 722 AGTAAGTAGTTGCCAGTATATAGTTTGGCMAAGTTGTTGTCATTTGCTACAGGCATCGT 663
QY 3001 GGTGTACGCTCGTGGTTGGTATGGCTTTCATTACAGTCCGGTCCCAACGATCAAGGCG 3060
Db 662 GGTGTACGCTCGTGGTTGGTATGGCTTTCATTACAGTCCGGTCCCAACGATCAAGGCG 603
QY 3061 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTCTCCGATCGT 3120
Db 602 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTCTCCGATCGT 543
QY 3121 TGTGCAAGTAGTTGGCGCAGGTGTTATCACTCATGTTTATGGCAGCAGCTGCAATAATC 3180
Db 542 TGTGCAAGTAGTTGGCGCAGGTGTTATCACTCATGTTTATGGCAGCAGCTGCAATAATC 483
QY 3181 TCTTACTGTCATGCTCCATCCGTAAGATGCTTTTCTGTGACTGTTAGTACTCAACCAAGTC 3240
Db 482 TCTTACTGTCATGCTCCGTAAGATGCTTTTCTGTGACTGTTAGTACTCAACCAAGTC 423
QY 3241 ATTCTGAGAATAGTGTATGCGCGCAGCGAGTTGCTCTTTCGGCGGCTCAATACGGGATAA 3300
Db 422 ATTCTGAGAATAGTGTATGCGCGCAGCGAGTTGCTCTTTCGGCGGCTCAATACGGGATAA 363
QY 3301 TACCGCGGCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAATAACGTTCTTCGGGGCG 3360
Db 362 TACCGCGGCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAATAACGTTCTTCGGGGCG 303

QY 3361 AAACTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCACTCGTGCACC 3420
Db 302 AAACTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCACTCGTGCACC 243
QY 3421 CAACTGATCTTTCAGCATCTTTTACTTTTCCAGAGCGTTTCTGGGTGAGCAAAACAGGAAG 3480
Db 242 CAACTGATCTTTCAGCATCTTTTACTTTTCCAGAGCGTTTCTGGGTGAGCAAAACAGGAAG 183
QY 3481 GCAAAATCCGCAAAAAGGGAATTAAGGCGACACGGAATGTTGAATACTCATACTCTT 3540
Db 182 GCAAAATCCGCAAAAAGGGAATTAAGGCGACACGGAATGTTGAATACTCATACTCTT 123
QY 3541 CCTTTTCAATATATTATTCGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
Db 122 CCTTTTCAATATATTATTCGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 53
QY 3601 TGAATGTATTTAGAAAAATAAAATAAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCC 3660
Db 62 TGAATGTATTTAGAAAAATAAAATAAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCC 3
QY 3661 AC 3662
Db 2 AC 1

RESULT 14
US-10-637-758-17/c
; Sequence 17, Application US/10637758
; Publication No. US20040180352A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/637,758
; PRIOR FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: US/10/280,913
; PRIOR FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Encodes cycle 3 GFP
US-10-637-758-17

Query Match 72.8%; Score 2664.8; DB 18; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;
QY 1 CTAAATTGTAAAGCGTTAATATTTTGTTAAAAATCGCGTTAAATTTTGTAAATCAGCTC 60
Db 3637 CTAAATTGTAAAGCGTTAATATTTTGTTAAAAATCGCGTTAAATTTTGTAAATCAGCTC 3578
QY 61 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATAAATCAAAAGATAGACCGA 120
Db 3577 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATAAATCAAAAGATAGACCGA 3518
QY 121 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAAGCTGGACTC 180
Db 3517 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAAGCTGGACTC 3458
QY 181 CAACGTCAAGGGCGAAAAAACCCGCTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240

QY 2401 AAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACACCCGCTGCTAGCGGTGGTTTTTT 2460
DB 1262 AAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACACCCGCTGCTAGCGGTGGTTTTTT 1203
QY 2461 TGTTCGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCTCTTTGATCTT 2520
DB 1202 TGTTCGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCTCTTTGATCTT 1143
QY 2521 TTCTACGGGTCTGACGCTCAGTGGAACGAAACCTCAGTTTAAGGATTTTGGTCATCAG 2580
DB 1142 TTCTACGGGTCTGACGCTCAGTGGAACGAAACCTCAGTTTAAGGATTTTGGTCATCAG 1083
QY 2581 ATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAATGAAATTTTAAATCAAT 2640
DB 1082 ATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAATGAAATTTTAAATCAAT 1023
QY 2641 CTAAGATATATAGTAAACCTTGTCTGACAGTTTACCAATGCTTAATCACTGAGGACCC 2700
DB 1022 CTAAGATATATAGTAAACCTTGTCTGACAGTTTACCAATGCTTAATCACTGAGGACCC 963
QY 2701 TATCTCAGGATCTGCTATTTGTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGAT 2760
DB 962 TATCTCAGGATCTGCTATTTGTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGAT 903
QY 2761 AACTACGATACGGAGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGCGAGACCC 2820
DB 902 AACTACGATACGGAGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGCGAGACCC 843
QY 2821 ACGTCAACGGCTCAGATTATACAGCAATAAACAGCAGCGGAGGCGGCGAGCGAG 2880
DB 842 ACGTCAACGGCTCAGATTATACAGCAATAAACAGCAGCGGAGGCGGCGAGCGAG 783
QY 2881 AAGTGTCTCGCACTTTATCGGCTCCATCCAGTCTATTAAATTTGTCGGGAGCTAG 2940
DB 782 AAGTGTCTCGCACTTTATCGGCTCCATCCAGTCTATTAAATTTGTCGGGAGCTAG 723
QY 2941 AGTAAGTAGTTCCGACGTTAATAGTTTGGCAACGTTGTGCCATTGTCTACAGGATCGT 3000
DB 722 AGTAAGTAGTTCCGACGTTAATAGTTTGGCAACGTTGTGCCATTGTCTACAGGATCGT 663
QY 3001 GGTGTACGCTCGTTCGTTGGTATGCTTCAATTCAGCTCCGGTTCCCAACGATCAAGGCG 3060
DB 662 GGTGTACGCTCGTTCGTTGGTATGCTTCAATTCAGCTCCGGTTCCCAACGATCAAGGCG 603
QY 3061 AGTTACATGATCCCACTGTTGTGCABAAAGCGGTAGCTCCTTCGGTCCCTCCGATCGT 3120
DB 602 AGTTACATGATCCCACTGTTGTGCABAAAGCGGTAGCTCCTTCGGTCCCTCCGATCGT 543
QY 3121 TGTTCAGAAAGTAAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATTC 3180
DB 542 TGTTCAGAAAGTAAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATTC 483
QY 3181 TCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTC 3240
DB 482 TCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTC 423
QY 3241 ATTCTGAGAAATAGTATGTCGGCGACCGAGTTGCTTTTCCCGCGGCTCAATACGGGATAA 3300
DB 422 ATTCTGAGAAATAGTATGTCGGCGACCGAGTTGCTTTTCCCGCGGCTCAATACGGGATAA 363
QY 3301 TACCGCGCACATACGAGAACTTTTAAAGTGCTCATCATTTGGAACACGTTCTTCGGGCGG 3360
DB 362 TACCGCGCACATACGAGAACTTTTAAAGTGCTCATCATTTGGAACACGTTCTTCGGGCGG 303
QY 3361 AAAACTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTGATTAACCACTCTGTGACCC 3420
DB 302 AAAACTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTGATTAACCACTCTGTGACCC 243
QY 3421 CAACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTCGGTCAGCAAAACAGGAAG 3480
DB 242 CAACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTCGGTCAGCAAAACAGGAAG 183

QY 3481 GCAAAATCCCAAAAAAGGAAATAAGGGCGACACGGAATGTTGAATATCTCATCTCTT 3540
DB 182 GCAAAATCCCAAAAAAGGAAATAAGGGCGACACGGAATGTTGAATATCTCATCTCTT 123
QY 3541 CCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGGGGATACATATT 3600
DB 122 CCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGGGGATACATATT 63
QY 3601 TGAATGTATTTAGAAAAATAAAACAAATAGGGGTTCCGGCGCACATTTCCCGAAAAAGTGCC 3660
DB 62 TGAATGTATTTAGAAAAATAAAACAAATAGGGGTTCCGGCGCACATTTCCCGAAAAAGTGCC 3
QY 3661 AC 3662
DB 2 AC 1

RESULT 15
US-10-066-390-3/c
; Sequence 3, Application US/10066390
; Publication No. US20020146732A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Lindbo, John A.
; APPLICANT: Fitzmaurice, Wayne P.
; TITLE OF INVENTION: A Method of Increasing Complementarity
; FILE REFERENCE: P-LG 4878
; CURRENT APPLICATION NUMBER: US/10/066,390
; CURRENT FILING DATE: 2002-02-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-066-390-3

Query Match 72.5%; Score 2656.4; DB 13; Length 3637;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;

QY 1 CTAATTCGTAAGCGTTAATATTGTTTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60
DB 3637 CTAATTCGTAAGCGTTAATATTGTTTAAATTCGCGTTAAATTTTGTAAATCAGCTC 3578
QY 61 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATATAATCAAAGAATAGACCGA 120
DB 3577 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATATAATCAAAGAATAGACCGA 3518
QY 121 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAAGATCCACTATTAAAGACGTGACCTC 180
DB 3517 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAAGATCCACTATTAAAGACGTGACCTC 3458
QY 181 CAACGTCAAAGCGGCAAAACCGCTCATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
DB 3457 CAACGTCAAAGCGGCAAAACCGCTCATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398
QY 241 CTAATCAAGTTTTTTGGGGTCCAGGTGCGTAAAGCACTAAATCGGAACCCCTAAAGGGAG 300
DB 3397 CTAATCAAGTTTTTTGGGGTCCAGGTGCGTAAAGCACTAAATCGGAACCCCTAAAGGGAG 3338
QY 301 CCCCAGATTTAGAGTTGACCGGGNAAGCCGCGAACCGTGGCGAGAAAGGAGGAGGAAGAA 360
DB 3337 CCCCAGATTTAGAGTTGACCGGGNAAGCCGCGAACCGTGGCGAGAAAGGAGGAGGAAGAA 3278
QY 361 AGCGAAAGGAGCGGCGCTAGGCGCTGCGCAAGTGTAGCGGTACGCTGCGGTAAACCAAC 420
DB 3277 AGCGAAAGGAGCGGCGCTAGGCGCTGCGCAAGTGTAGCGGTACGCTGCGGTAAACCAAC 3218
QY 421 CACACCCCGCGCTTAAATAGCGCGCTACAGGGCGCGTCCCAATTCGCCATTCAGGCTGCG 480

Db 3217 CACACCGCGCGCTTAATGCGCGCTACAGGGCGTCCCATTCGCCATTGAGGCTGCG 3158
Qy 481 CAACTGTTGGGAAGGCGATCGGTGCGGCTCTTTCGCTATTACGCCAGCTGCGGAAAGG 540
Db 3157 CAACTGTTGGGAAGGCGATCGGTGCGGCTCTTTCGCTATTACGCCAGCTGCGGAAAGG 3098
Qy 541 GGGATGTCGCAAGCGGATTAAGTTGGGTAAGCCAGGAGGTTTCCAGTCAGAGCTTG 600
Db 3097 GGGATGTCGCAAGCGGATTAAGTTGGGTAAGCCAGGAGGTTTCCAGTCAGAGCTTG 3038
Qy 601 TAAACGACGCGCAGTACGCGGCTCTGCTTCAATTCACGTTTGTGAACCGCTGAGGACGG 660
Db 3037 TAAACGACGCGCAGTACGCGGCTCTGCTTCAATTCACGTTTGTGAACCGCTGAGGACGG 2978
Qy 661 GCAGACTCGCGTGCAAAATGTTTTPACAGCGTGAATGGAGCAGATGAAGATGCTCGACAC 720
Db 2977 CCGCGTGGCGCGCTCTAGAACTAGTGGATCCCGCGGCTGCAGGAATTCCTATTGCT 2918
Qy 721 GCTGCAGAACGCGAGCTAGATTAAACCTAGAACGATTAATCATATTGTCAGCTACGTTAA 780
Db 2917 ATAGTTTCAATCCATGCCATGTAATCCAGCAGCTGTTTACAACTCAAGAGGACCATGT 2858
Qy 781 AGATAATC--ATGCGTAAATTTGACGATGGATCTGTAATACGACTCACTATAGGCGA 838
Db 2857 GGTCTCTCTTTTGGTGGATCTTTTGAAGGCGAGATTGTGTGACAGGTAATGGTGT 2798
Qy 839 ATTGGGTACCGGCGCCCCCTCGAGTGCAGCGTATCGATAAGCTTGTATCGAAATTCCT 898
Db 2797 CTGGTAAAGGACAGGCGCATGCCAAATGGAGTATTTTGTGATAATGGTCTGCTAGTT 2738
Qy 899 GCAGCGCGGGGATCCACTAGTTCATAGAGCGCGCCACCGCGGTGGAGCTCCAGCTTTT 958
Db 2737 GAAACGCTTCAATCTTCAATGTTGTCTAATTTTGAAGTTAACTTTTGAATTCATCTTTT 2678
Qy 959 GTTCCCTTTAGTGGGTTAATTTAGATCCATGCGTCAATTTTACGACACTCTTTCT 1018
Db 2677 GTTGTCTGCCAATGATATCAATTTGTGTA--GTTATAGTTGATTTCCAAATTTGTGTTCA 2619
Qy 1019 AGGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGCGCTCAGTCATCG 1078
Db 2618 AGAATGTTTCCATCTTCTTAAATCAATACCTTTTAACTCGATTC-----TATTAAACAA 2564
Qy 1079 CCCAAGCTGGCGCTATCTGGGATACAGTTTCGTGGGCTTTTTCGAGACACAGTCCGGAT 1138
Db 2563 GGGTATCACCTTCAAACCTTGAATTC-----AGCACGTGCTTGTAGTTCCTGTC 2515
Qy 1139 TGAAGCGCATGGAAGAGTTTGGCGAGGATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1198
Db 2514 ATCTTTGAAAAATATAGTTCTTTCTCTGTATACATAACCTTCGGGCTGCGCACTCTTGA 2455
Qy 1199 AGCGACGTTTACCATGATGATTTCGGGAAGGTGTGGCCATGCGCCTTTTAAACGCTGA 1258
Db 2454 GTCATGCGGTTTATATGATCTGGGTATCTTGAAGACATTGAAACCATATAGAAAGT 2395
Qy 1259 GTTCGTTACGGCCACTGGGATACAGTTTCGTGGGCTTTTTCGAGACACAGTCCGGAT 1318
Db 2394 AGTGACAGTGTGGCCATGGAACAGGTAGTTTTCAGTAGTGCAATTAATTTAAGGCT 2335
Qy 1319 GGTACGCGGACGCGATCAGCAACCCGAAATATACCGCGGACAGCGGAACTGCGTGC 1378
Db 2334 AAGTTTTC-----CGTATGTTGCATCACCTTCACCTCTCCACTGACACGAAATTTGTGC 2280
Qy 1379 CGGTGTGAGATTAATGACAGCGGTGCGGCTGGGATATTACGTACGCGGACGCGGTA 1438
Db 2279 CCAATTAACATCACCATCTAATTTCAACAGAAATTTGGACACTCCAGTGAAGTCTTCT 2220
Qy 1439 TCTGGCTGGATGCGCGAGAAATGGACATGGATACCCCGTGGATTTACCGCGGCGCGC 1498
Db 2219 CTTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAGGTTA-----ATTGCGGC 2165
Qy 1499 TTGGGCTAATCATGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCAATTTCA 1558

Db 2164 TTGGCGTAAATCATGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATCCA 2105
Qy 1559 CACAAATACGAGCGGAAAGCATAAAGTGTAAAGCCTGGGGTGCTTAATGAGTGAGCTAA 1618
Db 2104 CACAAATACGAGCGGAAAGCATAAAGTGTAAAGCCTGGGGTGCTTAATGAGTGAGCTAA 2045
Qy 1619 CTCACATTAATTTGGCTTGGCTCACCTGCGGCTTTTTCAGTCCGGAACCTGTGTCGCCAG 1678
Db 2044 CTCACATTAATTTGGCTTGGCTCACCTGCGGCTTTTTCAGTCCGGAACCTGTGTCGCCAG 1985
Qy 1679 CTGCAATTAATTTGGCCTGAGCGCGGAGAGGCGGTTTTCGTTATTTGGCGCTCTTCC 1738
Db 1984 CTGCAATTAATTTGGCCTGAGCGCGGAGAGGCGGTTTTCGTTATTTGGCGCTCTTCC 1925
Qy 1739 GCTTCTCTGCTCACTGACTCGCTGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1798
Db 1924 GCTTCTCTGCTCACTGACTCGCTGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1865
Qy 1799 CACTCAAAAGGCGGTAATACGTTATCCAGAAATCAGGGGATAACCGAGAAAGACATG 1858
Db 1864 CACTCAAAAGGCGGTAATACGTTATCCAGAAATCAGGGGATAACCGAGAAAGACATG 1805
Qy 1859 TGAGCAAAAGGCGGTAATACGTTATCCAGAAATCAGGGGATAACCGAGAAAGACATG 1918
Db 1804 TGAGCAAAAGGCGGTAATACGTTATCCAGAAATCAGGGGATAACCGAGAAAGACATG 1745
Qy 1919 CATAGCTTCCGCGCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGCGCA 1978
Db 1744 CATAGCTTCCGCGCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGCGCA 1685
Qy 1979 AACCGACAGGACTATAAGATACAGGCGTTTCCCGCTTGGAGCTCCCTCGTCCGCTCT 2038
Db 1684 AACCGACAGGACTATAAGATACAGGCGTTTCCCGCTTGGAGCTCCCTCGTCCGCTCT 1625
Qy 2039 CCGTTTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTCTCCTCTTCCGGAAGCGTG 2098
Db 1624 CCGTTTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTCTCCTCTTCCGGAAGCGTG 1565
Qy 2099 GCGCTTCTCATAGCTACGCTGTAGGTATCTCAGTTCCGTTAGGTGCTGCTCCGCTCAAG 2158
Db 1564 GCGCTTCTCATAGCTACGCTGTAGGTATCTCAGTTCCGTTAGGTGCTGCTCCGCTCAAG 1505
Qy 2159 CTGGGCTGTGTCAGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGCTAATAT 2218
Db 1504 CTGGGCTGTGTCAGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGCTAATAT 1445
Qy 2219 CGTCTTTAGTCCAAACCGGTAAGACAGACTTATCGCCACTGCGACGACCTGCTGTAAC 2278
Db 1444 CGTCTTTAGTCCAAACCGGTAAGACAGACTTATCGCCACTGCGACGACCTGCTGTAAC 1385
Qy 2279 AGGATTAGCAGAGCGAGGTATGTAGCGGTGCTCAGAGTCTTCTGAAGTGGTGGCTTAAC 2338
Db 1384 AGGATTAGCAGAGCGAGGTATGTAGCGGTGCTCAGAGTCTTCTGAAGTGGTGGCTTAAC 1325
Qy 2339 TACGGCTACATAGAAAGGACAGTATTTGGTATCTGGGCTCTGCTGAGGACAGTTACCTTC 2398
Db 1324 TACGGCTACATAGAAAGGACAGTATTTGGTATCTGGGCTCTGCTGAGGACAGTTACCTTC 1265
Qy 2399 GGAAGAGGTTGGTGTAGCTCTTGATCCGCAACCAACCCGCTGCTGAGCGTGGTGT 2458
Db 1264 GGAAGAGGTTGGTGTAGCTCTTGATCCGCAACCAACCCGCTGCTGAGCGTGGTGT 1205
Qy 2459 TTTGTTTTCGACAGCAGAGTATACGCGGAGAAAAAGGATCTCAAGAGATCCTTTGATC 2518
Db 1204 TTTGTTTTCGACAGCAGAGTATACGCGGAGAAAAAGGATCTCAAGAGATCCTTTGATC 1145
Qy 2519 TTTTCTACGGGTGTGACGCTCAGTGGAAACGAAACTCAAGTTAAGGATTTTGGTCATG 2578
Db 1144 TTTTCTACGGGTGTGACGCTCAGTGGAAACGAAACTCAAGTTAAGGATTTTGGTCATG 1085
Qy 2579 AGATTATCAAAAGGATCTTCACTAGTCTTCTTAAATTTAAATGAAAGTTTAAATCA 2638
Db 1084 AGATTATCAAAAGGATCTTCACTAGTCTTCTTAAATTTAAATGAAAGTTTAAATCA 1025

Search completed: June 6, 2005, 04:28:09
Job time : 2022.86 secs

```
QY 2639 ATCTAAAGTATATATGAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA 2698
Db 1024 ATCTAAAGTATATATGAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA 965
QY 2699 CCTATCTCAGCGATCTGCTCTATTTTCGTTTCATCCATAGATTGCTGACTCCCGTCTGTGTAG 2758
Db 964 CCTATCTCAGCGATCTGCTCTATTTTCGTTTCATCCATAGATTGCTGACTCCCGTCTGTGTAG 905
QY 2759 ATAACTAGGATACGGGAGGCTTACCATCTGCCCCAGTGTGCAATGATACCGCGAGAC 2818
Db 904 ATAACTAGGATACGGGAGGCTTACCATCTGCCCCAGTGTGCAATGATACCGCGAGAC 845
QY 2819 CCAGCTCACCGGCTCCAGATTTTATCAGCAATAAACACAGCCGAGCGGAGGCGCGAGCGC 2878
Db 844 CCAGCTCACCGGCTCCAGATTTTATCAGCAATAAACACAGCCGAGCGGAGGCGCGAGCGC 785
QY 2879 AGAAGTGTCTCTGCAACTTTTATTCGCTTCCTCCATCCAGTCTATTAATTGTTCGCGGAAGCT 2938
Db 784 AGAAGTGTCTCTGCAACTTTTATTCGCTTCCTCCATCCAGTCTATTAATTGTTCGCGGAAGCT 725
QY 2939 AGAGTAAAGTAGTTCGCCAGTTAAAGTTTTCGCAACGTTTTCGCAATGCTTACAGGCATC 2998
Db 724 AGAGTAAAGTAGTTCGCCAGTTAAAGTTTTCGCAACGTTTTCGCAATGCTTACAGGCATC 665
QY 2999 GTGGTGTCAACGCTCGTCTGTTTGGTATGGCTTCATTTCAGCTCCGGTTCCTCAACGATCAAGG 3058
Db 664 GTGGTGTCAACGCTCGTCTGTTTGGTATGGCTTCATTTCAGCTCCGGTTCCTCAACGATCAAGG 605
QY 3059 CGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGTTAGCTCCTTCGGTCTCTCCGATC 3118
Db 604 CGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGTTAGCTCCTTCGGTCTCTCCGATC 545
QY 3119 GTTCTCAGAACTAAGTTGGCGCGAGTGTATACCTCATGTTATGGCAGCACTGCATAAT 3178
Db 544 GTTCTCAGAACTAAGTTGGCGCGAGTGTATACCTCATGTTATGGCAGCACTGCATAAT 485
QY 3179 TCTCTTACTGTTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAG 3238
Db 484 TCTCTTACTGTTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAG 425
QY 3239 TCATTCTGAGAAATAGTGTATCGGCGACCCGAGTTGCTCTTCCCGGGGTCAATACGGGAT 3298
Db 424 TCATTCTGAGAAATAGTGTATCGGCGACCCGAGTTGCTCTTCCCGGGGTCAATACGGGAT 365
QY 3299 AATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATTTGGAAACGTTCTTCGGGG 3358
Db 364 AATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATTTGGAAACGTTCTTCGGGG 305
QY 3359 CGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGCA 3418
Db 304 CGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGCA 245
QY 3419 CCCAACTGATCTTCAGCATCTTTTATCTTTTACACGAGTTTCTGGGTGAGCAAAAACAGGA 3478
Db 244 CCCAACTGATCTTCAGCATCTTTTATCTTTTACACGAGTTTCTGGGTGAGCAAAAACAGGA 185
QY 3479 AGGCAAAATGCCGCAAAAAGGGAATAGGGGACACGGAATCTTGAATACTCATATCTC 3538
Db 184 AGGCAAAATGCCGCAAAAAGGGAATAGGGGACACGGAATCTTGAATACTCATATCTC 125
QY 3539 TTCCTTTTTCATTAATTATTGAAGCATTTTATCAGGGTTATTGTCTCATGAGCGGATACATA 3598
Db 124 TTCCTTTTTCATTAATTATTGAAGCATTTTATCAGGGTTATTGTCTCATGAGCGGATACATA 65
QY 3599 TTTGAAATGTTATTAGAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGGAAAAGTG 3658
Db 64 TTTGAAATGTTATTAGAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGGAAAAGTG 5
QY 3659 CCAC 3662
Db 4 CCAC 1
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 13:37:08 ; Search time 15351.2 Seconds
(without alignments)
11558.877 Million cell updates/sec

Title: US-10-826-523-41
Perfect score: 3662
Sequence: 1 ctaaattgtaagcgttaata.....atttccccgaaaagtgcac 3662

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sta.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query %	Length	DB ID	Description
C 1	2664.8	72.8	3637	6	AX592968	AX592968 Sequence
C 2	2656.4	72.5	3637	6	AX592967	AX592967 Sequence
C 3	2650.6	72.4	3626	12	Y09373	Y09373 Artificial
C 4	2644.8	72.2	3624	12	XXU43957	XXU43957 Expression
5	2639	72.1	3681	6	A20702	A20702 pKSEL5 DNA
6	2639	72.1	3681	6	I15651	I15651 Sequence 6
7	2623.6	71.6	5314	6	A20700	A20700 pMTVI DNA s
8	2623.6	71.6	5314	6	I15649	I15649 Sequence 4
C 9	2610.8	71.3	3552	12	AF171227	AF171227 Shuttle v
C 10	2609.4	71.3	4001	6	AX306330	AX306330 Sequence
C 11	2558.8	69.9	5540	12	AF179627	AF179627 Expressio
C 12	2550.6	69.7	3633	12	ASFGRN2	Y09374 Artificial
C 13	2517.8	68.8	3918	6	CQ768840	CQ768840 Sequence
C 14	2517.8	68.8	3918	6	CQ802124	CQ802124 Sequence
C 15	2517.8	68.8	3918	6	CQ873225	CQ873225 Sequence
C 16	2517.8	68.8	3918	6	CQ876882	CQ876882 Sequence
C 17	2514.6	68.7	3886	12	XXU35129	U35129 Plasmid pBS
C 18	2485.8	67.9	4141	12	AX190524	AX190524 Cloning v
C 19	2484.6	67.8	3774	6	AX3063326	AX3063326 Sequence

20	2474.2	67.6	5277	6	BD244458	BD244458 Genome an
21	2469.4	67.4	5277	6	AR340565	AR340565 Sequence
22	2451.8	67.0	3928	6	AX591027	AX591027 Sequence
23	2451.8	67.0	3928	6	AX591180	AX591180 Sequence
24	2451.8	67.0	3928	6	AX717603	AX717603 Sequence
25	2449.8	66.9	7933	12	AF359244	AF359244 Integrati
26	2448.8	66.9	3956	6	A84776	A84776 Sequence 21
27	2448.8	66.9	3956	6	AR300179	AR300179 Sequence
28	2448.8	66.9	3956	6	BD058066	BD058066 Method fo
29	2446.8	66.8	3715	6	AX306328	AX306328 Sequence
30	2434	66.5	3970	12	CVE403983	AJ403983 Cloning v
31	2420.6	66.1	3274	12	XXU43956	U43956 Expression
C 32	2414.6	65.9	3978	12	CUV41513	U41513 3' terminal
C 33	2401	65.6	3733	12	AF146395	AF146395 Cloning v
C 34	2396	65.4	3307	12	SYNPFABCX	D50399 Unidentifie
C 35	2382.6	65.1	4591	12	SYNPEZ218A	M74186 Cloning vec
C 36	2380.4	65.0	6624	12	CVU37458	U37458 Yeast CUP1
C 37	2372.2	64.8	4239	12	AVE401048	AJ401048 Artificial
38	2359	64.4	5781	12	CVU37457	U37457 Yeast CUP1
39	2357.4	64.4	4051	12	XXU35126	U35126 Plasmid pBS
C 40	2351	64.2	5275	12	CVU37456	U37456 Yeast CUP1
41	2349.8	64.2	3230	12	AY663843	AY663843 Cloning v
42	2340.2	63.9	3230	12	AY663844	AY663844 Cloning v
C 43	2333.4	63.7	5436	12	AY303166	AY303166 CFP Integ
C 44	2332.8	63.7	5436	12	AY303167	AY303167 YFP Integ
45	2328.6	63.6	3204	12	AY663845	AY663845 Cloning v

ALIGNMENTS

RESULT 1
LOCUS AX592968/c 3637 bp DNA linear PAT 27-JAN-2003
DEFINITION Sequence 4 from Patent WO02079468.
ACCESSION AX592968
VERSION AX592968.1 GI:27950812
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Padgett,H.S., Fitzmaurice,W.P. and Lindo,J.A.
TITLE A method of increasing complementarity in a heteroduplex
JOURNAL Patent: WO 02079468-A 4 10-OCT-2002;
Large Scale Biology Corporation (US)
FEATURES
source
1..3637
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

ORIGIN	Query Match	Best Local Similarity	Mismatches	Conservative	Score	Pred. No. 0;	DB 6;	Length	3637;
QY	1	CTAAATTGTAAGCGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTC	60						
DB	3637	CTAAATTGTAAGCGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTC	3578						
QY	61	ATTTTAAACCAATAGGCGGAAATCGGAAATCCCTTATAATCAAAAGAAATAGACCGA	120						
DB	3577	ATTTTAAACCAATAGGCGGAAATCGGAAATCCCTTATAATCAAAAGAAATAGACCGA	3518						
QY	121	GATAGGTTGAGTGTGTTCCAGTTTGGACCAAGTCCACTATTAAAGAACTGGACTC	180						
DB	3517	GATAGGTTGAGTGTGTTCCAGTTTGGACCAAGTCCACTATTAAAGAACTGGACTC	3458						
QY	181	CAACGTCAAAGGCGGAAACCGTCTATCAGGCGGATGCCCACTACGTGAACCATCACC	240						
DB	3457	CAACGTCAAAGGCGGAAACCGTCTATCAGGCGGATGCCCACTACGTGAACCATCACC	3398						

QY 241 CTAATCAAGTTTTTGGGCTGAGGTGCGTAAAGCACTAAATCGGAACCCCTAAAGGAG 300
DB |||||
3397 CTAATCAAGTTTTTGGGCTGAGGTGCGTAAAGCACTAAATCGGAACCCCTAAAGGAG 3338
QY 301 CCCCAGATTGAGCTTGAACGGGGAAGACCGCGGAACGTGGCGAGAAAGGAAGGAAGAA 360
DB |||||
3337 CCCCAGATTGAGCTTGAACGGGGAAGACCGCGGAACGTGGCGAGAAAGGAAGGAAGAA 3278
QY 361 AGCGAAAGAGCGGCGCTAGGCGCTGCAAGTGTAGCGGTCAGCGTCGCGTAACCAAC 420
DB |||||
3277 AGCGAAAGAGCGGCGCTAGGCGCTGCAAGTGTAGCGGTCAGCGTCGCGTAACCAAC 3218
QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGCGCGTCCCATTCGCCATTGAGGCTGG 480
DB |||||
3217 CACACCGCGCGCTTAATGCGCGCTACAGGCGCGTCCCATTCGCCATTGAGGCTGG 3158
QY 481 CAACTGTTGGGAAGGCGGATCGGTGGGCTCTTTCGCTATTATACGCACTGGCGAAAGG 540
DB |||||
3157 CAACTGTTGGGAAGGCGGATCGGTGGGCTCTTTCGCTATTATACGCACTGGCGAAAGG 3098
QY 541 GGGATGCTGCAAGCGGATTAAGTTGGTAAAGCGGCTTCCAGTCAAGGCTGG 600
DB |||||
3097 GGGATGCTGCAAGCGGATTAAGTTGGTAAAGCGGCTTCCAGTCAAGGCTGG 3038
QY 601 TAAACGACGCGCAGTGAGCGCGCTCGTTCAATTCACGTTTTTGAACCCGTGGAGGACGG 660
DB |||||
3037 TAAACGACGCGCAGTGAGCGCGCTCAATACGACTCACTATAGGGCGAATTGGAGCTCA 2978
QY 661 GCAGACTGCGGTGCAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
DB |||||
2977 CCGCGTGGCGG-----CCGCTCTAGAACTAGTGGATCCG--CGG 2939
QY 721 GCTGCAGAACCGCAGCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTAGTTAA 780
DB |||||
2938 GCTGCAGAAATCTTATTGTTGATGTTTCATCGCATGTGTAATCCAGCAGCAGTTA 2879
QY 781 AGATAATCATGCTGTAATTTGACGATGGGATCTGTAAATACGACTCACTATAGGGCGGAAT 840
DB |||||
2878 CAAACTCAAGAGGACCATGTGTTGCTGTTAAAGGACAGGGCCATCGCCAAATGGAGTATTT 2819
QY 841 TGGGTACCGGCGCCCTCGAGGTGCGAGGTATCGATAAGTTGATATCGAATTCCTGCG 900
DB |||||
2818 GTGTCGACAGGTAAATGTTGTTGCTGTTAAAGGACAGGGCCATCGCCAAATGGAGTATTT 2759
QY 901 AGCCCGGGGATCCACTAGTTCTAGAGCGCGCCACCGCGTGGAGTCCAGCTTTTGT 960
DB |||||
2758 GTTGATATGCTGCTAGTTGAACGGATCCATCTTCAATGTTGTGGCAATTTGAAGT 2699
QY 961 TCCCTTTAGTGAAGGTTAATTAGATCCCATGCGTCAATTTTACGCAAGCTATCTTTCTAG 1020
DB |||||
2698 TAGCTTTGATTCCATCTTTTGTGCTGCGTGATGTATACATTGTGTGAG-TTATAG 2640
QY 1021 GGTAACTAGCTGATCAGGATCATATCGTGGGTCTTTTTCGGCTCAGTCAATCGCC 1080
DB |||||
2639 TTGTACTCGAGTTTGTGTCGGAATGTTTCCATCTCTTTTAAATCAATACCTTTTAAAC 2580
QY 1081 CAAAGTGGCGCTATCTGGCATCGGGAGGAAGACCGTGCCCTTTTCCGCGAGGTTG 1140
DB |||||
2579 TCGATACGATTAACAGGGTATCACCTTCAAACTTGACCTTACGACGCGCTTTGTAGTTC 2520
QY 1141 AAGCGGCATGGAAGAGTTTGGCGAGGATGACTGCTGTGATGACGTTGAGCGTGAAGAAAC 1200
DB |||||
2519 CCGTCATCTTTGAAAGATATAGTGGCTTCTGTACATAACCTTCGGGATGGCACTCTTG 2460
QY 1201 GCACGTTTACGATGATTCGGGAAGGTGTGGCCATGACGCGCTTTAAACGTGAATCTGT 1260
DB |||||
2459 AAAAAGTCATGCGGTTTCATATGATCCGGATAACCGGGAAGGATTTGAACACCAATAGAG 2400
QY 1261 TCGTTTCAGGCCACCTGGGATACAGTTGCTCGCGCTTTTTCGGGACACAGTTTCCGGATGG 1320
DB |||||
2399 AAAAGTAGTACAGTGTGGCCATCGAACAGGTAGTTTTCAGTAGTAGTCAATAAATTTA 2340
QY 1321 TCAGCCGGAAGCGCATACGAAACCCGAAATAATACCGGCGACAGCCGGAACCTGCGGTGGC 1380

DB 2339 AGGGTAAGCTTTTCGTATGTAGCATCACTTACCTCTCCACTGACAGAAATTTGTGC 2280
QY 1381 GTGTCCAGATTAAATGACAGCGGTGGCGGTATATACGTGACGAGGAGGGGTATC 1440
DB 2279 CCAATTAACATCACTCAATTAATTAACAAGAAATGGGACAACCTCCAGTGAAGATTTCTTCT 2220
QY 1441 CTGGCTGATGCCGAGAAATGGACATGATACCCCGTGAGTTACCCGCGCGGCGCTT 1500
DB 2219 C---CTTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTAAATTTGCGGCTT 2163
QY 1501 GCGCTAATCATGGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATTTCCACA 1560
DB 2162 GCGCTAATCATGGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATTTCCACA 2103
QY 1561 CAACTATGAGCCGGAAGCATAAAGTGTAAAGCTTGGGTGCTTAATGATGAGCTAACT 1620
DB 2102 CAACTATGAGCCGGAAGCATAAAGTGTAAAGCTTGGGTGCTTAATGATGAGCTAACT 2043
QY 1621 CACTTAATTTGGTTGCGCTCACTGCGCTTCCAGTCCGGAACCTGTGCGGAGCT 1680
DB 2042 CACTTAATTTGGTTGCGCTCACTGCGCTTCCAGTCCGGAACCTGTGCGGAGCT 1983
QY 1681 GCATTAATGAATTCGCCAAACGCGCGGAGAGCGGTTTGCCTATTTGGGCGCTCTTCCGC 1740
DB 1982 GCATTAATGAATTCGCCAAACGCGCGGAGAGCGGTTTGCCTATTTGGGCGCTCTTCCGC 1923
QY 1741 TTCCTCGCTCACTGACTGCTGCGCTCGCTGCTTCCAGTCCGCGAGCGGATATCAGTCA 1800
DB 1922 TTCCTCGCTCACTGACTGCTGCGCTCGCTGCTTCCAGTCCGCGAGCGGATATCAGTCA 1863
QY 1801 CTCGAGGCGGTAATACGCTTATCCACAGAAATCAGGGGATTAACGAGGAAAGAAATGTG 1860
DB 1862 CTCGAGGCGGTAATACGCTTATCCACAGAAATCAGGGGATTAACGAGGAAAGAAATGTG 1803
QY 1861 AGCAAAAGGCGCAGCAAAAGGCGGAGAACCGTAAAGGCGGCTTGTGGCGTTTTCCTCA 1920
DB 1802 AGCAAAAGGCGCAGCAAAAGGCGGAGAACCGTAAAGGCGGCTTGTGGCGTTTTCCTCA 1743
QY 1921 TAGGCTCGCGCCCTGAGCAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 1980
DB 1742 TAGGCTCGCGCCCTGAGCAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 1683
QY 1981 CCCGACAGGACTATAAAGATACAGGCGTTTCCCTCTGGAAGCTCCCTGTCGCTCTCC 2040
DB 1682 CCCGACAGGACTATAAAGATACAGGCGTTTCCCTCTGGAAGCTCCCTGTCGCTCTCC 1623
QY 2041 TGTTCGACCCCTGCGCTTACCGGATACGTGCGCTTTTCTCCCTTCGGGAGGCTGCG 2100
DB 1622 TGTTCGACCCCTGCGCTTACCGGATACGTGCGCTTTTCTCCCTTCGGGAGGCTGCG 1563
QY 2101 GCTTTCATAGCTCACGCTGATGATATCAGTTTCGGGTAGTGTGCTTCCCTCCAAGCT 2160
DB 1562 GCTTTCATAGCTCACGCTGATGATATCAGTTTCGGGTAGTGTGCTTCCCTCCAAGCT 1503
QY 2161 GGGCTGTGTGACGAAACCCCGCTTTCAGCCGCGCTGCGCTTATCCGCTTAACCTATCG 2220
DB 1502 GGGCTGTGTGACGAAACCCCGCTTTCAGCCGCGCTGCGCTTATCCGCTTAACCTATCG 1443
QY 2221 TCTTGAAGTCCAAACCGGTTAAGCAGCATTTATCGGCTGCGGAGCGGCTGCTGTAACG 2280
DB 1442 TCTTGAAGTCCAAACCGGTTAAGCAGCATTTATCGGCTGCGGAGCGGCTGCTGTAACG 1383
QY 2281 GATTAGCAGAGCGAGTATGAGGCGGTGCTACAGATTTCTTGAAGTGGTGGCTTAACTA 2340
DB 1382 GATTAGCAGAGCGAGTATGAGGCGGTGCTACAGATTTCTTGAAGTGGTGGCTTAACTA 1323
QY 2341 CCGCTTACACTAGAAGCAGATTTTGGTATCTGCGCTCTGCTGTAAGCAGGTTACCTCGG 2400
DB 1322 CCGCTTACACTAGAAGCAGATTTTGGTATCTGCGCTCTGCTGTAAGCAGGTTACCTCGG 1263
QY 2401 AAAAGAGTTGGTAGCTTTCATTCGGGCAACAAACCCGCTGCTGAGCGGTGTTTTT 2460

Db 1262 AAAAAAGTTGGTAGCTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGTTTTT 1203
QY 2461 TGTTCAGACGACAGATTACGGCGAGAAAAAGGATCTCAAGAAGATCCTTTTGATCTT 2520
Db 1202 TGTTCGAAGCAGCAGATTACGGCGAGAAAAAGGATCTCAAGAAGATCCTTTTGATCTT 1143
QY 2521 TTCTACGGGTCTGACGCTCAGTGGGAACGAAACTCAGTTAAGGATTTTGTCTATGAG 2580
Db 1142 TTCTACGGGTCTGACGCTCAGTGGGAACGAAACTCAGTTAAGGATTTTGTCTATGAG 1083
QY 2581 ATTATCAAAAAGGATCTTCACTCTAGATCTTTTAAATTTAAATGAAGTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGGATCTTCACTCTAGATCTTTTAAATTTAAATGAAGTTTAAATCAAT 1023
QY 2641 CTAAAGTATATATAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGCACC 2700
Db 1022 CTAAAGTATATATAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGCACC 963
QY 2701 TATCTCAGGATCTGTCTATTTCGTTTCATCCATAGTTGCTGACTCCCGTCTGTGTAGAT 2760
Db 962 TATCTCAGGATCTGTCTATTTCGTTTCATCCATAGTTGCTGACTCCCGTCTGTGTAGAT 903
QY 2761 AACTAGATACGGAGGGCTTACCATCTGSCCCAGTCTGCAATGATACCGGAGACCC 2820
Db 902 AACTAGATACGGAGGGCTTACCATCTGSCCCAGTCTGCAATGATACCGGAGACCC 843
QY 2821 ACCTCACCGGCTCCAGATTTATCAGCAATAAACACGACCGAGCGGAGCGGAG 2880
Db 842 ACCTCACCGGCTCCAGATTTATCAGCAATAAACACGACCGAGCGGAGCGGAGCGAG 783
QY 2881 AGTGTCTCTGCACTTTATCCGCTCTCATCTAGTCTATTAATGTTGCGGGAAGCTAG 2940
Db 782 AGTGTCTCTGCACTTTATCCGCTCTCATCTAGTCTATTAATGTTGCGGGAAGCTAG 723
QY 2941 AGTAACTAGTTCGCCAGTTAATAGTTTGGCAACGTTTGTCCATGCTACAGGATCTGT 3000
Db 722 AGTAACTAGTTCGCCAGTTAATAGTTTGGCAACGTTTGTCCATGCTACAGGATCTGT 663
QY 3001 GGTGTCAACGCTCGTCTGTTGGTATGCTTCATTCAGCTCCGGTTCCTCAACGATCAAGGCG 3060
Db 662 GGTGTCAACGCTCGTCTGTTGGTATGCTTCATTCAGCTCCGGTTCCTCAACGATCAAGGCG 603
QY 3061 AGTTACATGATCCCAATGTTGTGCAAAAAGGCGTTAGCTCTTCGGTCTCCGATCTGT 3120
Db 602 AGTTACATGATCCCAATGTTGTGCAAAAAGGCGTTAGCTCTTCGGTCTCCGATCTGT 543
QY 3121 TGTCAAGTAAAGTTCGCCGCGAGTTATCACTCATGTTATGCGAGCACTGCATATTC 3180
Db 542 TGTCAAGTAAAGTTCGCCGCGAGTTATCACTCATGTTATGCGAGCACTGCATATTC 483
QY 3181 TCTTACTGTCAATGCCATCCGTAAAGATGCTTTTCTGTGACTGGTGGTACTCAACCAAGTC 3240
Db 482 TCTTACTGTCAATGCCATCCGTAAAGATGCTTTTCTGTGACTGGTGGTACTCAACCAAGTC 423
QY 3241 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCTCAATACGGGATAA 3300
Db 422 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCTCAATACGGGATAA 363
QY 3301 TACCGGCGCACATAGCAGAACTTTAAAGTCTCATCATTTGGAAGACGTTCTTCGGGGCG 3360
Db 362 TACCGGCGCACATAGCAGAACTTTAAAGTCTCATCATTTGGAAGACGTTCTTCGGGGCG 303
QY 3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCAGTGTAAACCACTCGTGACCC 3420
Db 302 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCAGTGTAAACCACTCGTGACCC 243
QY 3421 CAACTGATCTTACGATCTTTTACTTTTCAACGAGTTTCTGGGTGAGCAAAAACAGGAAG 3480
Db 242 CAACTGATCTTACGATCTTTTACTTTTCAACGAGTTTCTGGGTGAGCAAAAACAGGAAG 183
QY 3481 GCNAATTCGCGCAAAAAGGGAATAGGGGACACCGAATGTTGAATCTACTACTCTT 3540
Db 182 GCNAATTCGCGCAAAAAGGGAATAGGGGACACCGAATGTTGAATCTACTACTCTT 123

QY 3541 CCTTTTCAATATATTGAAGCAATTTATCAGGGTTATTTGCTCATGAGCGGATACATATT 3600
Db 122 CCTTTTCAATATATTGAAGCAATTTATCAGGGTTATTTGCTCATGAGCGGATACATATT 63
QY 3601 TGAATGTATTTAGAAAAATAACAAATAGGGTTCCGCGCACATTTCCCGGAAAAAGTGCC 3660
Db 62 TGAATGTATTTAGAAAAATAACAAATAGGGTTCCGCGCACATTTCCCGGAAAAAGTGCC 3
QY 3661 AC 3662
Db 2 AC 1
RESULT 2
LOCUS AX592967/c 3637 bp DNA linear PAT 27-JAN-2003
DEFINITION Sequence 3 from Patent WO02079468.
ACCESSION AX592967
VERSION AX592967.1 GI:27950811
KEYWORDS synthetic construct
SOURCE other sequences; artificial sequences.
ORGANISM 1
REFERENCE 1
AUTHORS Padgett,H.S., Fitzmaurice,W.P. and Lindo,J.A.
TITLE A method of increasing complementarity in a heteroduplex
JOURNAL Patent: WO 02079468-A 3 10-OCT-2002;
Large Scale Biology Corporation (US)
FEATURES
source 1. .3637
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 72.5%; Score 2656.4; DB 6; Length 3637;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;
QY 1 CTAAATTTGAAGCGTTAATAATTTTGTAAATTCGGGTTAAATTTTGTAAATCAGCTC 60
Db 3637 CTAAATTTGAAGCGTTAATAATTTTGTAAATTCGGGTTAAATTTTGTAAATCAGCTC 3578
QY 61 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 120
Db 3577 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 3518
QY 121 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
Db 3517 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458
QY 181 CAACGTCAAAGGGCGGAAACCGTCTATCAGGGCGATGGCCCACTAGTGAACCATCACC 240
Db 3457 CAACGTCAAAGGGCGGAAACCGTCTATCAGGGCGATGGCCCACTAGTGAACCATCACC 3398
QY 241 CTAATCAAGTTTTTTGGGGTTCGAGGTGCGGTAAAGCACTAAATCGGAACCTTAAAGGGAG 300
Db 3397 CTAATCAAGTTTTTTGGGGTTCGAGGTGCGGTAAAGCACTAAATCGGAACCTTAAAGGGAG 3338
QY 301 CCCCCGATTTAGAGCTTGACGCGGAAAGCCGGCGAAACGTTGGCGGAGAAAGGAAGAA 360
Db 3337 CCCCCGATTTAGAGCTTGACGCGGAAAGCCGGCGAAACGTTGGCGGAGAAAGGAAGAA 3278
QY 361 ACGGAAGAGCGGGCGCTAGGGCGCTAGGGCTGAGCGGTAGCGGTCACTGCGCGTAAACAC 420
Db 3277 ACGGAAGAGCGGGCGCTAGGGCGCTAGGGCTGAGCGGTAGCGGTCACTGCGCGTAAACAC 3218
QY 421 CACACCCGCGCGCTTTAATGCGCGCTACAGGCGCGTCCCATTCGCCATTTCAGGCTGCG 480
Db 3217 CACACCCGCGCGCTTTAATGCGCGCTACAGGCGCGTCCCATTCGCCATTTCAGGCTGCG 3158
QY 481 CAACTGTTGGGAAGGGCGGATTCGGTTCGGGCGCTCTTCGCTATTACGCCAGCTGCGGAAGG 540

Db 3157 CAACTGTTGGGAAGGCGCATCGGTGGGCGCTCTTTCGCTATTATACGCCAGCTGGCGAAAGG 3098
QY 541 GGGATGTCGCTCAAGCGGCAATTAAGTTGGGTACGCCAGGTTTTCCCACTCAGTCAAGCGTTG 600
Db 3097 GGGATGTCGCTCAAGCGGCAATTAAGTTGGGTACGCCAGGTTTTCCCACTCAGTCAAGCGTTG 3038
QY 601 TAAAAACGCGCCAGTCAGCGCGCTCGTTCAATTCACGTTTTTGAACCCCGTGGAGGAGG 660
Db 3037 TAAAAACGCGCCAGTCAGCGCGCTCGTTCAATTCACGTTTTTGAACCCCGTGGAGGAGG 2978
QY 661 GCAGACTCGCGTGCAAAATGTTTTACAGCGTGATGAGAGAGATGAAGATGCTCGACAC 720
Db 2977 CGCGGTGGCGCGCTCTAGAACTAGTGATCCCGCGGCTGCAGGAATCTTATTGTT 2918
QY 721 GCTGCAGAACGCGAGCTAGATTAAACCTTAGAAGATTAATCATATGTTGACGTACGTTAA 780
Db 2917 ATAGTTCAATCCATGCCATGTGTAATCCCCAGCAGCTGTTACAAACTCAAGAGGACATGT 2858
QY 781 AGATAATC - ATGCGTAAATTTGACGCATGGGATCTGTAATACGACTCACTATAGGGGA 838
Db 2857 GGTCTCTCTTTTCGTTGGATCTTTTCGAAAGGCGAGATTGTGTGACAGGTAATGTTGT 2798
QY 839 ATTGGGTACCGGGCCCCCTCCAGGTCCAGGTCCAGGTATCGATAAGCTTTGATATCGAATTCCT 898
Db 2797 CTGGTAAAGGACAGGGCCATCGCCAAATTGGAGTATTTGTTGATAATGGTCTGCTAGTT 2738
QY 899 GCAGCCGGGGATCCACTAGTTCTTAGAGCGGCCCCACCGGGTGGAGCTCCAGCTTTT 958
Db 2737 GAACGCTTCCATCTTCAATGTTGTGCTAAATTTTGAAGTTAACTTTGATTCGATCTTTT 2678
QY 959 GTTCCCTTTAGTGGGTTAAATTAGATCCATGCGTCAATTTTACGACACTATCTTTCT 1018
Db 2677 GTTTGCTGCCATGATGATACATTTGTGTGA - GTTATAGTTGATTTCCAAATTGTTGCCA 2619
QY 1019 AGGGTTAATCTAGTGCATCAGGATCATATGTCGGGTCTTTTTTCCGGCTCAGTCACTG 1078
Db 2618 AGAATGTTTCCATCTCTTTAAATCAATACCTTTTAACTCGATTC - - - - - TATTAAACA 2564
QY 1079 CCCAGCTGGGCTATCTGGGATCGGGGAGGAGAACCGCGTCTTTCCGCGAGT 1138
Db 2563 GGGTATACCTTCAACTTGACTTC - - - - - AGCACGTGCTTTGATGTTCCCGTC 2515
QY 1139 TGAAGCGCATGGAAGAGTTTGGCGAGGATGACTGCTGCTCAATGACGTTGAGCGGAA 1198
Db 2514 ATCTTTGAAAATATAGTTCTTCTCTGTACATAACCTTCGGGCATGGCACTCTTGAANA 2455
QY 1199 AGCAGCTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCGTTTAAACGGTGAAC 1258
Db 2454 GTCATGCGGTTTCAATGATCTGGGTATCTTGAAGAAGCATTTGAACACCATAGAGAAAGT 2395
QY 1259 GTTCGTTCAGGCCACTGGGATACAGTTTCGTCGGGCTTTTCGGGACACAGTCCGGAT 1318
Db 2394 AGTGACAAGTGTGGCCATGGAAACAGGTAGTTTTCAGTAGTGCAAAATAATTTAAGGT 2335
QY 1319 GGTGAGCCGAGCGCATCAGCAACCGCAACATACCGCGCAGCGGAACTGCGCTGC 1378
Db 2334 AAGTTTTC - - - - - CGTATGTTGCATCACCTTACCCCTCTCCACTGACAGAAATTTGTGC 2280
QY 1379 CGGTGTGAGATTAATGACAGCGGTGCGCGCTGGGATATTAACGTACGAGGACGGGTA 1438
Db 2279 CCATTAACATCACCATCTAATTAACAAGAAATGGGACAACTCCAGTGAAGAAAGTTCTTCT 2220
QY 1439 TCCTGGCTGGATGCCGAGAAATGACATGGATACCCGTGAGTTACCGCGGGCGGC 1498
Db 2219 CCTTTACTCATCGTACCCAGCTTTTGTTCCTTTTAGTGAGGGTTA - - - - - ATTCGCGC 2165
QY 1499 TTGGCGTAATCATGTCTAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTTCA 1558
Db 2164 TTGGCGTAATCATGTCTAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTTCA 2105
QY 1559 CACAACTACGAGCGGAGCATAAAGCTTGAAGCTTGGGTCCTTAATGATGAGCTAA 1618

Db 2104 CACAACTACGAGCGGAGCATAAAGTTGTAAGCCTGGGGTGCTTAATGATGAGCTAA 2045
QY 1619 CTCACATTAATTTGCGTTGCGCTCACTGCGCGCTTTCAGTTCGGGAAACCTGTCGTGCCAG 1678
Db 2044 CTCACATTAATTTGCGTTGCGCTCACTGCGCGCTTTCAGTTCGGGAAACCTGTCGTGCCAG 1985
QY 1679 CTGCATTAATGAATCGGCCAACGCGCGGGAGAGCGGTTTGCCTATTGGGCGCTCTTCC 1738
Db 1984 CTGCATTAATGAATCGGCCAACGCGCGGGAGAGCGGTTTGCCTATTGGGCGCTCTTCC 1925
QY 1739 GCTTCTCGCTCAGTCACTGCTGCGCTCGTTCGTTGCTGCGGCGGAGCGGTATCAGCT 1798
Db 1924 GCTTCTCGCTCAGTCACTGCTGCGCTCGTTCGTTGCTGCGGCGAGCGGTATCAGCT 1865
QY 1799 CACTCAAAAGGCGTAAATACGGTTATCCACAGATCAGGGGATAACGAGGAAAGAACATG 1858
Db 1864 CACTCAAAAGGCGTAAATACGGTTATCCACAGATCAGGGGATAACGAGGAAAGAACATG 1805
QY 1859 TGAGCAAAAGGCGCAGCAAAAGGCCAGGAACCGTAAAAAGGCGCGTTGCTGCGGTTTTTC 1918
Db 1804 TGAGCAAAAGGCGCAGCAAAAGGCCAGGAACCGTAAAAAGGCGCGTTGCTGCGGTTTTTC 1745
QY 1919 CATAGGCTCGCGCCCTTCGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGGA 1978
Db 1744 CATAGGCTCGCGCCCTTCGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGGA 1685
QY 1979 AACCCGACAGGACTATAAAGATACCAAGCGTTTTCCCTCGGAAGCTCCCTCGTGGCTCT 2038
Db 1684 AACCCGACAGGACTATAAAGATACCAAGCGTTTTCCCTCGGAAGCTCCCTCGTGGCTCT 1625
QY 2039 CCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTG 2098
Db 1624 CCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTG 1565
QY 2099 GCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTTCGTCGCTCCAG 2158
Db 1564 GCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTTCGTCGCTCCAG 1505
QY 2159 CTGGGCTGTGTGACGAAACCCCGCTTCAAGCCGACCGCTGCGCTTATPCCGGTAACAT 2218
Db 1504 CTGGGCTGTGTGACGAAACCCCGCTTCAAGCCGACCGCTGCGCTTATPCCGGTAACAT 1445
QY 2219 CGTCTTGTAGTCCAAACCCGGTAAGACACGACTTATCGCACTGGCAGAGCCACTGGTAAC 2278
Db 1444 CGTCTTGTAGTCCAAACCCGGTAAGACACGACTTATCGCACTGGCAGAGCCACTGGTAAC 1385
QY 2279 AGGATTACAGAGCAGGTATGTAGCGGTGCTACAGAGTTCCTTCAAGTGTGGGCTTAAC 2338
Db 1384 AGGATTACAGAGCAGGTATGTAGCGGTGCTACAGAGTTCCTTCAAGTGTGGGCTTAAC 1325
QY 2339 TACGGCTACACTAGAAAGGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTC 2398
Db 1324 TACGGCTACACTAGAAAGGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTC 1265
QY 2399 GGAAGAAAGAGTTGTGTAGCTTTGATCCGCGAAACAAACACCGCTGGTAGCGGTGTTTT 2458
Db 1264 GGAAGAAAGAGTTGTGTAGCTTTGATCCGCGAAACAAACACCGCTGGTAGCGGTGTTTT 1205
QY 2459 TTTGTTTCCAGCAGCAGATTTACGCGCAGAAAAAAGGATCTCAAGAGATCCTTTGATC 2518
Db 1204 TTTGTTTCCAGCAGCAGATTTACGCGCAGAAAAAAGGATCTCAAGAGATCCTTTGATC 1145
QY 2519 TTTTCTCAGCGGTCTGACGCTCAGTGGAAACGAAAACTCACGTTTAAGGGATTTTGGTCATG 2578
Db 1144 TTTTCTCAGCGGTCTGACGCTCAGTGGAAACGAAAACTCACGTTTAAGGGATTTTGGTCATG 1085
QY 2579 AGATTATCAAAAGGATCTTCACTAGATCTTTTAAATTTAAAAATGAAGTTTTAAATCA 2638
Db 1084 AGATTATCAAAAGGATCTTCACTAGATCTTTTAAATTTAAAAATGAAGTTTTAAATCA 1025
QY 2639 ATCTAAAGTATATGAGTAACTTTGCTGACAGTTACCAATGCTTAATCAGTGAGGCA 2698
Db 1024 ATCTAAAGTATATGAGTAACTTTGCTGACAGTTACCAATGCTTAATCAGTGAGGCA 965

```
QY 2699 CTTATCTCAGCGATCTGTCTATTTCGTTTCATTCATAGTTGCCTGACTCCCGTCGTGTAG 2758
DB 964 CTTATCTCAGCGATCTGTCTATTTCGTTTCATTCATAGTTGCCTGACTCCCGTCGTGTAG 905
QY 2759 ATAACTACGATACGGAGGCTTACCATCTGGCCCGAGTCTGCTCAATGATACCGCGAGAC 2818
DB 904 ATAACTACGATACGGAGGCTTACCATCTGGCCCGAGTCTGCTCAATGATACCGCGAGAC 845
QY 2819 CCAGCTCACCCTCTCAGATTATCAGCAATAAACACAGCCAGCCGGAAGGCGCGAGCGC 2878
DB 844 CCAGCTCACCCTCTCAGATTATCAGCAATAAACACAGCCAGCCGGAAGGCGCGAGCGC 785
QY 2879 AGAAGTGTCTCTCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGCGGAGCT 2938
DB 784 AGAAGTGTCTCTCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGCGGAGCT 725
QY 2939 AGAGTAAGTAGTTCCGAGTTAATAGTTTCGCAAGCTTGTTCGCCAATTTGTCGCGGAGCT 2998
DB 724 AGAGTAAGTAGTTCCGAGTTAATAGTTTCGCAAGCTTGTTCGCCAATTTGTCGCGGAGCT 665
QY 2999 GTGGTGTACGCTCTGCTGTGTATGGCTTCAATCAGCTCCGGTTCCTCAACGATCAAG 3058
DB 664 GTGGTGTACGCTCTGCTGTGTATGGCTTCAATCAGCTCCGGTTCCTCAACGATCAAG 605
QY 3059 CGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGTCTCTTCGGTCTCTCGATC 3118
DB 604 CGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGTCTCTTCGGTCTCTCGATC 545
QY 3119 GTTGTGAGAGTAAGTTGGCGGAGTGTATCACTCATGTTATGCGACGACTGCATAAT 3178
DB 544 GTTGTGAGAGTAAGTTGGCGGAGTGTATCACTCATGTTATGCGACGACTGCATAAT 485
QY 3179 TCTCTTACTGTCTATGCTATCGGATGCTTTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3238
DB 484 TCTCTTACTGTCTATGCTATCGGATGCTTTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 425
QY 3239 TCATTCGAGAAATGATGATGCGGCGACCGAGTTGCTTCTTGGCCGCGTCAATACGGAT 3298
DB 424 TCATTCGAGAAATGATGATGCGGCGACCGAGTTGCTTCTTGGCCGCGTCAATACGGAT 365
QY 3299 ATACCGCGCCACATAGCAACTTTTAAAGTCTCATCATTCGAAACGTTCTTCGGG 3358
DB 364 ATACCGCGCCACATAGCAACTTTTAAAGTCTCATCATTCGAAACGTTCTTCGGG 305
QY 3359 CGAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA 3418
DB 304 CGAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA 245
QY 3419 CCCAACTGATCTTCAGATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAAACAGGA 3478
DB 244 CCCAACTGATCTTCAGATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAAACAGGA 185
QY 3479 AGCAAAATGCCCAAAAAGGAATAAGCGGACACCGGAATGTTGAATCTCATCTC 3538
DB 184 AGCAAAATGCCCAAAAAGGAATAAGCGGACACCGGAATGTTGAATCTCATCTC 125
QY 3539 TTCTTTTCAATATTATTAAGCACTTTATCAGGGTTATTGTTCTCATGAGCGGATACATA 3598
DB 124 TTCTTTTCAATATTATTAAGCACTTTATCAGGGTTATTGTTCTCATGAGCGGATACATA 65
QY 3599 TTTGAATGTTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCATATTTCCCGAAAAGTG 3658
DB 64 TTTGAATGTTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCATATTTCCCGAAAAGTG 5
QY 3659 CCAC 3662
DB 4 CCAC 1
```

RESULT 3

ASPGREEN1/c

LOCUS

ASPGREEN1

3626 bp

DNA

linear

SYN 15-MAY-1998

```
DEFINITION Artificial sequences, plasmid vector pGreen-1.
ACCESSION Y09373
VERSION Y09373.1
KEYWORDS GI:1684624
          blatem gene; lacZ-phoC fusion gene; multiple cloning site; T3
          promoter.
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
          other sequences; artificial sequences.
AUTHORS Thaller,M.C., Berlutti,F., Schippa,S., Selan,L. and Rossolini,G.M.
TITLE Bacterial acid phosphatase gene fusions useful as targets for
JOURNAL cloning-dependent insertional inactivation
MEDLINE Biotechnol. Prog. 14 (2), 241-247 (1998)
PUBMED 98215688
REFERENCE 2 (bases 1 to 3626)
          Rossolini,G.M.
AUTHORS Direct Submission
TITLE Submitted (11-NOV-1996) G.M. Rossolini, Dip. Biologia
JOURNAL Moleculare-Sez., Microbiologia, Univ. di Siena, via Laterina N.8,
          I- 53100 Siena, ITALY
COMMENT Related sequences: X64444 and X52328.
FEATURES
    Location/Qualifiers
        1..3626
            /organism="synthetic construct"
            /mol_type="other DNA"
            /db_xref="taxon:32630"
            /plasmid="pGreen-1"
            129..989
                /gene="blatem"
            129..989
                /gene="lacZ-phoC fusion"
            129..989
                /codon_start=1
            129..989
                /protein_id="CAA70547.1"
            129..989
                /translation="MSIQHPRVALIPFFAFCPLPVFAHPETLVKVKDAEDQLGARVGY
            129..989
                /translation="ISDLSNGKILSFPRFPMSTFVLLCGAVLSRISDAGQEQLGRRIRHYSQNDLVE
            129..989
                /translation="YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLTTIGGPKELTAFLEHMGDHFRL
            129..989
                /translation="DWPELNEAIPIEDERTTTPVAMATTLKLLTGLLTLASRQLLDMKADKVAQPL
            129..989
                /translation="LRSALPAGNFIADKSGAGRGSGRIIAALGPDGKPSRIVIVYTTGSGATMDERNQIA
            129..989
                /translation="EIGASLIKHV"
            129..989
                /note="pUC M13/Reverse Sequencing primer recognition site"
            129..989
                /gene="lacZ-phoC fusion"
            129..989
                /codon_start=1
            129..989
                /protein_id="CAA70548.1"
            129..989
                /translation="MTMTTPSAQLTLTKGNKWSSTVAALALVLPDPAIPAGNDATTK
            129..989
                /translation="PDLVYKNEQALISLKLPPPEVSGIQFLNDQAMYEKGRMLRNLERGRKQAOADADLA
            129..989
                /translation="AGCVATASGAGPYITEKDSPELYKLLNMLIEDAGDLATRSKAEHYMRIRPFAYGT
            129..989
                /translation="ETCNTDKQKLSYSGTSIGWATLVIAEVNPNPANDRAKLEQYQLGSGRVICG
            129..989
                /translation="YHMQSDVDAARIVGSAANAATLHSDPAFQALAKAKQEPACKSQK"
            129..989
                /gene="lacZ-phoC fusion"
            129..989
                /note="T3 promoter"
primer_bind
    129..989
        /note="pUC M13/Reverse Sequencing primer recognition site"
    129..989
        /gene="lacZ-phoC fusion"
    129..989
        /codon_start=1
    129..989
        /protein_id="CAA70548.1"
    129..989
        /translation="MTMTTPSAQLTLTKGNKWSSTVAALALVLPDPAIPAGNDATTK
    129..989
        /translation="PDLVYKNEQALISLKLPPPEVSGIQFLNDQAMYEKGRMLRNLERGRKQAOADADLA
    129..989
        /translation="AGCVATASGAGPYITEKDSPELYKLLNMLIEDAGDLATRSKAEHYMRIRPFAYGT
    129..989
        /translation="ETCNTDKQKLSYSGTSIGWATLVIAEVNPNPANDRAKLEQYQLGSGRVICG
    129..989
        /translation="YHMQSDVDAARIVGSAANAATLHSDPAFQALAKAKQEPACKSQK"
    129..989
        /gene="lacZ-phoC fusion"
    129..989
        /note="T3 promoter"
promoter
    129..989
        /note="T3 promoter"
ORIGIN
    Query Match 72.4%; Score 2650.6; DB 12; Length 3626;
    Best Local Similarity 84.0%; Pred. No. 0;
    Matches 3079; Conservative 0; Mismatches 544; Indels 42; Gaps 6;
    QY 1 CTAATTGTAAGCGTTAATATTTTGTAAATTCGGTTAAATTTTGTAAATTTTGTAAATCAGTTC 60
    DB 3626 CTAATTGTAAGCGTTAATATTTTGTAAATTCGGTTAAATTTTGTAAATTTTGTAAATCAGTTC 3567
    QY 61 ATTTTAAACCAATAGCGCGAAATCGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 120
    DB 3566 ATTTTAAACCAATAGCGCGAAATCGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 3507
```

QY 121 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAAAGTGGACTC 180
Db |||||
QY 3506 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAAAGTGGACTC 3447
Db |||||
QY 181 CAAAGTCAAAAGGCGGAAACCGTCTATCAGGGCGATGCCCACTACGTGGAACCAATCAC 240
Db |||||
QY 3446 CAAAGTCAAAAGGCGGAAACCGTCTATCAGGGCGATGCCCACTACGTGGAACCAATCAC 3387
QY |||||
Db |||||
QY 241 CTAATCAAGTTTGGGTCGAGGTCCGTAAGCACTAAATCGGAACCCCTAAAGGGAG 300
Db |||||
QY 3386 CTAATCAAGTTTGGGTCGAGGTCCGTAAGCACTAAATCGGAACCCCTAAAGGGAG 3327
QY |||||
Db |||||
QY 301 CCCCCGATTAGAGCTTGACGGGGAAAGCCGGCGAAGCTGGCGAGAAAGAAAGGGAAGAA 360
Db |||||
QY 3326 CCCCCGATTAGAGCTTGACGGGGAAAGCCGGCGAAGCTGGCGAGAAAGAAAGGGAAGAA 3267
QY |||||
Db |||||
QY 361 AGCGAAAGAGCGGGCGCTAGGGCGTGGCAAGTGTAGCGGTACGCTGCGGTAAACCA 420
Db |||||
QY 3266 AGCGAAAGAGCGGGCGCTAGGGCGTGGCAAGTGTAGCGGTACGCTGCGGTAAACCA 3207
QY |||||
Db |||||
QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCCAATTCGCCAATTCAGGCTGCG 480
Db |||||
QY 3206 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCCAATTCGCCAATTCAGGCTGCG 3147
QY |||||
Db |||||
QY 481 CAACTGTTGGGAAGGCGGATCGGTGCGGCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
Db |||||
QY 3146 CAACTGTTGGGAAGGCGGATCGGTGCGGCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3087
QY |||||
Db |||||
QY 541 GGGATGTGTCGAAGCGGATTAAGTTGGGTAAACGCCAGAGGTTTCCAGTACGAGGTG 600
Db |||||
QY 3086 GGGATGTGTCGAAGCGGATTAAGTTGGGTAAACGCCAGAGGTTTCCAGTACGAGGTG 3027
QY |||||
Db |||||
QY 601 TAAACGCGGCGAGTACGCGGCTCGTTTCATTACAGTTTGAACCGTCGGAGGACGG 660
Db |||||
QY 3026 TAAACGCGGCGAGTACGCGGCTCGTTTCATTACAGTTTGAACCGTCGGAGGACGG 2967
QY |||||
Db |||||
QY 661 GCAGACTCCGCGTGCAAAATGTGTTTACAGCGTGATGGAGCAGATGAAGATGCTCGAC 720
Db |||||
QY 2966 GCCCCCTCGAGTTCGAGGTATCGATAAGCTTATCTGTGATTTTGTGCAAAATCC 2907
QY |||||
Db |||||
QY 721 GTCGAGAACACGCGCTAGATAACCCCTAGAAAGATAATATATGTGACGTACGTTAA 780
Db |||||
QY 2906 TGTGTTGGCTTCGCTAACTGCGCTGAATGCGGATCGGAATGTAAATGTCGCGACAGCG 2847
QY |||||
Db |||||
QY 781 AGATAATCATGCGTAAATTTGACGATGGGATCTGTATACGACTCACTATAGGCGAAT 840
Db |||||
QY 2846 GCTGAACC-----GACAACTCCGCGCGCATCCACATCACTCTGCGAGTATAGCCGA 2794
QY |||||
Db |||||
QY 841 TGGGTACCGGGGCCCCCTCGAGGTGCGAGGTATCGATAAGCTTGATATCGAATTCCTGCG 900
Db |||||
QY 2793 AATCACC CGGCTCTGTCCGAGCTGATAACCCGTTCCAGAATCGCATCCTGATTTGCCGG 2734
QY |||||
Db |||||
QY 901 AGCCGGGGGATCCACTAGTTCTAGAGCGGCGGCAACCGCGGTGAGCTCCAGCTTTTGT 960
Db |||||
QY 2733 GTTC----ACTTCCGCGACACTAATGCGGTTGCGCGAGCCGATAGACGTATGACCTGACG 2678
QY |||||
Db |||||
QY 961 TCCCTTTAGTAGGGGTTAATTAGATCCCATCGTCAATTTTACGCGAGACTATCTTTCTAG 1020
Db |||||
QY 2677 GGTAAAGATCCGTTGGTGGAGAGTTTTCCTGA-TCCTTGGTATTACAGGTTTCTGTGCGG 2619
QY |||||
Db |||||
QY 1021 GGTAAATCTAGCTGATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCAATCGCC 1080
Db |||||
QY 2618 TAAACGCAACCGACCGATGCGATGTAAATGTTCTTTGGCGGAGGGGTGCAAGGTCA 2559
QY |||||
Db |||||
QY 1081 CAACTGGCGCTATCTGGGCATCTGGGGAGGAAGAACCGCGTCCCTTTCCCGGAGGTTG 1140
Db |||||
QY 2558 CCGGCATCCTCAATCATATTGTCAGCAGTTTATACAGCTCCGAGAGTCTTTTTCGGTT 2499
QY |||||
Db |||||
QY 1141 AAGCGGCATGGAAGAGTTTTCGAGGATGACTGCTGCTGCAATTCAGCTTGAGCGGAAC 1200
Db |||||
QY 2498 ATCGGATAGCCGATGCCCCCTGAAATGTCGGTTGCCACCCCTCGCGCCAGGTACGA 2439

QY 1201 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCA---CGCCTTTAAACGGTGAAC 1257
Db |||||
QY 2438 TCTGCTGTGCTGTTTTCGCGCTCGGTATTGCGCAGCATACGCGCTTCTCATACATT 2379
Db |||||
QY 1258 TGTTCGTTCAGGCCACCTGGGATACAGTTGTCGCGGCTTTTCGGGACACAGTTTCCGA 1317
Db |||||
QY 2378 GCTGATCATTTTAAAACTGAATACTGCGGCTTCCGGCGGTGGCGGTAAACAGTTTCAGG 2319
QY |||||
Db |||||
QY 1318 TGTTCAGCCGGAAGCGCATCAGCAACCCGAAACAAATACCGCGCACAGCGCGAACTGCGGTG 1377
Db |||||
QY 2318 CTGTGATAGCTGTTCATTTTTCAGATAATAATAAT-----CCGCG 2277
QY |||||
Db |||||
QY 1378 CCGGTGTCAGATTAAATGACAGCGGTGCGGCTCGGATATTACGTACGAGGACGGT 1437
Db |||||
QY 2276 TTGGTGGTGGCATCGTTGCTGCGGGATCGCAGGATCCACTAGTTCTAGAGCGCGCGC 2217
QY |||||
Db |||||
QY 1438 ATCTGCGTGGATGCGGAGAAATGACATGATATACCGGTGAGTTACCCGCGGCGCGG 1497
Db |||||
QY 2216 ACCGCGTGGAGCTCCAGCTTTTGTTCCTTTAGTGAAGGTTAAATTT-----GCGCG 2166
QY |||||
Db |||||
QY 1498 CTTGCGGTAATCATGCTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAATTC 1557
Db |||||
QY 2165 CTTGCGGTAAATCATGCTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAATTC 2106
QY |||||
Db |||||
QY 1558 ACACAACATACAGCGCGGAAGCATAAAGTGTAAAGCCTGGGTGCTTAATGAGTGAGCTA 1617
Db |||||
QY 2105 ACACAACATACAGCGCGGAAGCATAAAGTGTAAAGCCTGGGTGCTTAATGAGTGAGCTA 2046
QY |||||
Db |||||
QY 1618 ACTCATTAATTGCGTTCGCTCACTGCGGCTTCCAGTCGGGAAACCTGCTGCGCA 1677
Db |||||
QY 2045 ACTCATTAATTGCGTTCGCTCACTGCGGCTTCCAGTCGGGAAACCTGCTGCGCA 1986
QY |||||
Db |||||
QY 1678 GCTGATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTTGGCGCTCTTC 1737
Db |||||
QY 1985 GCTGATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTTGGCGCTCTTC 1926
QY |||||
Db |||||
QY 1738 CGCTTCTCGTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1797
Db |||||
QY 1925 CGCTTCTCGTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1866
QY |||||
Db |||||
QY 1798 TCACCTCAAGGCGGTAAATACGTTTATCCACAGAAATCAGGGATACGAGGAAAGAAAT 1857
Db |||||
QY 1865 TCACCTCAAGGCGGTAAATACGTTTATCCACAGAAATCAGGGATACGAGGAAAGAAAT 1806
QY |||||
Db |||||
QY 1858 GTGACAAAAGGCCAGCAAAAAGGCCAGGAAACCGTAAAAAGGCCGCTTGTGCTGCTGCTT 1917
Db |||||
QY 1805 GTGACAAAAGGCCAGCAAAAAGGCCAGGAAACCGTAAAAAGGCCGCTTGTGCTGCTGCTT 1746
QY |||||
Db |||||
QY 1918 CCATAGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 1977
Db |||||
QY 1745 CCATAGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 1686
QY |||||
Db |||||
QY 1978 AAACCGGACGAGCTATAAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTGCTGCGCTC 2037
Db |||||
QY 1695 AAACCGGACGAGCTATAAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTGCTGCGCTC 1626
QY |||||
Db |||||
QY 2038 TCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCTTCTCCCTTCGGGAAGCGT 2097
Db |||||
QY 1625 TCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCTTCTCCCTTCGGGAAGCGT 1566
QY |||||
Db |||||
QY 2098 GCGCTTTCTCATAGCTCACTGCTGAGTATCTCAGTTTCGGGTAGGTGCTTCCGCTCAA 2157
Db |||||
QY 1565 GCGCTTTCTCATAGCTCACTGCTGAGGTATCTCAGTTTCGGGTAGGTGCTTCCGCTCAA 1506
QY |||||
Db |||||
QY 2158 GCTGGGCTGTGACGAAACCCCGCTCAGCCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2217
Db |||||
QY 1505 GCTGGGCTGTGACGAAACCCCGCTCAGCCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1446
QY |||||
Db |||||
QY 2218 TCGTCTGAGTCCAAACCCCGGTAAAGACAGCTTATCGGCACCTGGCAGCAGCACTGGTAA 2277
Db |||||
QY 1445 TCGTCTGAGTCCAAACCCCGGTAAAGACAGCTTATCGGCACCTGGCAGCAGCACTGGTAA 1386
QY |||||
Db |||||
QY 2278 CAGGATTAAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTTGAAGTGTGCGCTAA 2337

Db 1385 CAGATTAGCAGAGGATGATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCATA 1326
QY 2338 CTACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT 2397
Db 1325 CTACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT 1266
QY 2398 CGGAAAAGAGTTGGTAGCTCTTGTATCCGGCAACAAACACACCGCTGGTAGCGGTGTTT 2457
Db 1265 CGGAAAAGAGTTGGTAGCTCTTGTATCCGGCAACAAACACACCGCTGGTAGCGGTGTTT 1206
QY 2458 TTTTGTGTCAGCAGCAGATAGCGCGCAGAAAAGAGATCTCAAGAAGATCCTTTGAT 2517
Db 1205 TTTTGTGTCAGCAGCAGATAGCGCGCAGAAAAGAGATCTCAAGAAGATCCTTTGAT 1146
QY 2518 CTTTCTACGGGTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGGAATTTTGGTCA 2577
Db 1145 CTTTCTACGGGTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGGAATTTTGGTCA 1086
QY 2578 GAGATTATCAAAAAGGATCTTACCTAGATCCTTTTAAATTAATAAATGAAGTTTAAATC 2637
Db 1085 GAGATTATCAAAAAGGATCTTACCTAGATCCTTTTAAATTAATAAATGAAGTTTAAATC 1026
QY 2638 AATCTAAAGTATATAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGC 2697
Db 1025 AATCTAAAGTATATAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGC 966
QY 2698 ACCTATCTCAGCGATCTGTATTTGGTTCATCCATAGTTGCTGACTCCCGCTCGTGA 2757
Db 965 ACCTATCTCAGCGATCTGTATTTGGTTCATCCATAGTTGCTGACTCCCGCTCGTGA 906
QY 2758 GATAACTACGATACGGAGGCTTACATCTGCCCCAGTGTGCAATGATACCGGAGA 2817
Db 905 GATAACTACGATACGGAGGCTTACATCTGCCCCAGTGTGCAATGATACCGGAGA 846
QY 2818 CCACGCTCAGCGCTCCAGATTATACGAAATACCAACGCGGAGGCGGAGCG 2877
Db 845 CCACGCTCAGCGCTCCAGATTATACGAAATACCAACGCGGAGGCGGAGCG 786
QY 2878 CAGAAGTGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAGC 2937
Db 785 CAGAAGTGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAGC 726
QY 2938 TAGAGTAAGTAGTTGCGCAGTTAATAGTTTGGCGCAACGTTTGGTCCATGCTACAGCAT 2997
Db 725 TAGAGTAAGTAGTTGCGCAGTTAATAGTTTGGCGCAACGTTTGGTCCATGCTACAGCAT 666
QY 2998 CGTGTGTACGCTCGTCTGTTGGTATGCTTCAATCAGCTCCGTTCCCAACGATCAAG 3057
Db 665 CGTGTGTACGCTCGTCTGTTGGTATGCTTCAATCAGCTCCGTTCCCAACGATCAAG 606
QY 3058 GCAGATTACATGATCCCAATGTTGTGCAAAAAGCGGTTAGTCTCTTGGTCTCCGAT 3117
Db 605 GCAGATTACATGATCCCAATGTTGTGCAAAAAGCGGTTAGTCTCTTGGTCTCCGAT 546
QY 3118 CGTTGTACAGTAAGTTGCGCAGTGTATCACTCATGTTTATGCGACACTGCATAA 3177
Db 545 CGTTGTACAGTAAGTTGCGCAGTGTATCACTCATGTTTATGCGACACTGCATAA 486
QY 3178 TTCTCTTACTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGGTGGTACTCAACCAA 3237
Db 485 TTCTCTTACTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGGTGGTACTCAACCAA 426
QY 3238 GTCAATCTGAGATAGTTGATCGCGCAGCGAGTTGCTTTGCGCGGCTCAATACGGGA 3297
Db 425 GTCAATCTGAGATAGTTGATCGCGCAGCGAGTTGCTTTGCGCGGCTCAATACGGGA 366
QY 3298 TAATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCATTTGGAAGAGTTCTTCGCG 3357
Db 365 TAATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCATTTGGAAGAGTTCTTCGCG 306
QY 3358 GCGAAAACCTCTAAGGATCTTACCGCTGTTGAGATCAGTTTCGATGTAACCCACTCGTGC 3417

Db 305 GCGAAAACCTCTAAGGATCTTACCGCTGTTGAGATCAGTTTCGATGTAACCCACTCGTGC 246
QY 3418 ACCAACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGG 3477
Db 245 ACCAACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGG 186
QY 3478 AAGGCAAAATCGCGCAAAAAGGGAATAGGGCGACACGGAATGTTGAATACTCATACT 3537
Db 185 AAGGCAAAATCGCGCAAAAAGGGAATAGGGCGACACGGAATGTTGAATACTCATACT 126
QY 3538 CTTTCTTTTCAATATTTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACAT 3597
Db 125 CTTTCTTTTCAATATTTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACAT 66
QY 3598 ATTTGAATGTTATTAGAAAATAAACAATAGGGTTCGCGCACATTTTCCCGAAAAGT 3657
Db 65 ATTTGAATGTTATTAGAAAATAAACAATAGGGTTCGCGCACATTTTCCCGAAAAGT 6
QY 3658 GCCAC 3662
Db 5 GCCAC 1

RESULT 4
XXU43957
LOCUS
DEFINITION Expression vector pB10T-MRz, complete cds.
ACCESSION U43957
VERSION U43957.1 GI:4097019
KEYWORDS
SOURCE Expression vector pB10T-MRz
ORGANISM Expression vector pB10T-MRz
REFERENCE 1 (bases 1 to 3624)
AUTHORS Ruiz, J., Wu, C.H., Ito, Y. and Wu, G.Y.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1995) Juan Ruiz, Division of
Gastroenterology-Hepatology, University of Connecticut Health
Center, Farmington, CT 06030, USA
COMMENT In vivo expression vector for the synthesis of a ribozyme against
the core region of the Hepatitis B virus pregenomic RNA. It does
not recognize the viral core protein but the core region of the
viral RNA. pB10T-MRz contains a tandem of ten units of the
ribozyme, separated between them by the target sequence recognized
by the ribozyme, cloned at the XbaI-HindIII site of pBluescript SK
(+).

FEATURES
Location/Qualifiers
1..3624
/organism="Expression vector pB10T-MRz"
/mol_type="genomic DNA"
/db_xref="taxon:45892"
complement(657..1425)
/note="multiple cloning site"
complement(721..766)
/product="hammerhead ribozyme"
complement(791..836)
/product="hammerhead ribozyme"
complement(861..906)
/product="hammerhead ribozyme"
complement(931..976)
/product="hammerhead ribozyme"
complement(1001..1046)
/product="hammerhead ribozyme"
complement(1071..1116)
/product="hammerhead ribozyme"
complement(1141..1186)
/product="hammerhead ribozyme"
complement(1211..1256)
/product="hammerhead ribozyme"
complement(1281..1326)
/product="hammerhead ribozyme"
complement(1351..1396)
/product="hammerhead ribozyme"

promoter	complement (1482..1604)	
gene	/note="Lac-2"	
	complement (2639..3499)	
CDS	/gene="ampR"	
	complement (2639..3499)	
	/gene="ampR"	
	/function="ampicillin resistance marker"	
	/codon_start=1	
	/transl_table=11	
	/product="beta-lactamase"	
	/protein_id="AADI0137.1"	
	/db_xref="GI:4097020"	
	/translation="MSIQHFRVALIPFFAFLCPVPAHPETLVKVKDAEDQLGARVGY	
	IELDLSNGKILLESFPERFPMWSFKVLLCGAVLSRFDAGQQLGRIHYSONDLVE	
	YSPVTEKHLTDGMTVRELCSAAITWSDNTAANLLLTITGGPKELTAFILHNMGDHVTRL	
	DWEPELNEAIENDSERDTPMPVAMATLRLKLLTGLLTSLASRQQLIDNMEADKVAGPL	
	LRSLALPWFIADKSGAGSRGIIAALGPDGKFSRIYVIYITGSQATMDERNRQIIA	
	EIGASLIKHW"	
ORIGIN		
Query Match	72.2%; Score 2644.8; DB 12; Length 3624;	
Best Local Similarity	84.3%; Pred. No. 0;	
Matches 3096; Conservative	0; Mismatches 512; Indels 64; Gaps 8;	
QY	1 CTAATTTGTAAGCGTTAAATATTGTTTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60	
DB	4 CTAATTTGTAAGCGTTAAATATTGTTTAAATTCGCGTTAAATTTTGTAAATCAGCTC 63	
QY	61 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAATCAAAAGATAGACCGA 120	
DB	64 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAATCAAAAGATAGACCGA 123	
QY	121 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAAAGTGGACTC 180	
DB	124 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAAAGTGGACTC 183	
QY	181 CAACGTCAAAGGGCGAAACCGCTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240	
DB	184 CAACGTCAAAGGGCGAAACCGCTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 243	
QY	241 CTAATCAAGTTTTTGGGGTCGAGTGCCTGTAAGCACTAAATCGGAACCCCTAAAGGGAG 300	
DB	244 CTAATCAAGTTTTTGGGGTCGAGTGCCTGTAAGCACTAAATCGGAACCCCTAAAGGGAG 303	
QY	301 CCCCAGATTAGACTTACGGGGAAGCCGCGCAACGTGGCGAGAAAGGAAGGAAGAA 360	
DB	304 CCCCAGATTAGACTTACGGGGAAGCCGCGCAACGTGGCGAGAAAGGAAGGAAGAA 363	
QY	361 AGCGAAAGGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCAAC 420	
DB	364 AGCGAAAGGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCAAC 423	
QY	421 CACACCGCGCGCTTAATCGCGCGTACAGGGCGGTCCCATTCGCCATTCAGGCTGGG 480	
DB	424 CACACCGCGCGCTTAATCGCGCGTACAGGGCGGTCCCATTCGCCATTCAGGCTGGG 483	
QY	481 CAACTGTTGGGAAGGCGAGTCGTCGGGCTCTTCGCTATTATACGCAAGCTCGCGAAAGG 540	
DB	484 CAACTGTTGGGAAGGCGAGTCGTCGGGCTCTTCGCTATTATACGCAAGCTCGCGAAAGG 543	
QY	541 GGGATGTGCTCAAGGCGAATTAAGTTGGGTAAACGCCAGGGTTTTCCCAATCACGACGTTG 600	
DB	544 GGGATGTGCTCAAGGCGAATTAAGTTGGGTAAACGCCAGGGTTTTCCCAATCACGACGTTG 603	
QY	601 TAAACGACGGCCAGTG-----AGGGCGCTCGTTTCATTACGTTTTTGAACCCGTGGA 654	
DB	604 TAAACGACGGCCAGTGAAATTGTAATACGACTCACTATAGGGCGAATTGGGTACCGGGCC 663	
QY	655 GGACGGGACAGACTCGCGGTGCAAAATGTGTTTTACAGCGGTGATGGAGCAGATCAAGATGCT 714	
DB	664 CCCCCTCGAGGTCGAGGGTATCGATAAGCTTGAGACTAGTTCAGGAGAGACTCTGGATCCAA 723	
QY	715 CGACACGCTGCAGAACACACGACGTAGATTAAACCCCTAGAAAGATATATATTGTGACCGTA 774	

DB	724 CCTTAGAGTTTCGTCTCACGGACTCATCAGTCCT---GAGCATTTGTCTAGTTTCAGGAGA 780	
QY	775 CGTTAAAGATATCATCGTGTAAATTTGACGATGGATCTGTAAATACGACTCACTATAGG 834	
DB	781 CTCTGGATCCAAGCCTTAGAGTTTCGTCTCACGGACTCATCAGTCCTGAGCAATTGTCTTA 840	
QY	835 GCGAATTGGGTACCGGGCCCCCTCGAGGTCGACGGGTATCGATAGCTTGATATCGAAT 894	
DB	841 GTTCAGGAGACTCTCGATCCAAAGCCTTAGAGTTTCGTCTCACGGA-----886	
QY	895 TCCTCGACCGCGGGGATCCACTAGTTCTAGAGCGCGCCACCGCGGTGGAGTCCACG 954	
DB	887 CTCATCAGTCCTGAGCAATTGTCTAGTTCCAGGAGA-----C 921	
QY	955 TTTTGTTCCTTTAGTAGGAGTTAATTAGATCCCATCGTCAATTTTACGCAGACTATCT 1014	
DB	922 TCTGATCAAGCCTTAGAGTTTCGTCTCACGGACTCATCAGTCTCAGGCAATTGTCTAG 981	
QY	1015 TTCTAGGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTC 1074	
DB	982 TTCAGGAG-----ACTCTGGATCCAAGCCTTAGAGTTTCGTCTCACGGACTCATCA 1033	
QY	1075 ATCGCCCCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAACCCGTCCTTTTCCCGG 1134	
DB	1034 GTCCTGAGCAATTGTCTAGTTCAGGAGACTCTGGATCCAAAGCCTTAGAGTTTCGTCTCAC 1093	
QY	1135 AGTTTGAAGCGCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCATTTGACGTTGAGC 1194	
DB	1094 GGACTCATCAGTCTCAGGACTTTGTCTAGTTCAGGAGACT-CTGGATCCAAGCCTTAGAGT 1152	
QY	1195 GAAAAACGACGTTTACCATGATGATTCCGGAAGGTGTGCCCATGACGCTTTTAAACGGTG 1254	
DB	1153 TTCGTCTCTCAGGACTCATCAGTCTCTGAGCATTTGTCTAGTTTCAGGAGACTCTGGATCCA 1212	
QY	1255 AACTGTTCTGTCAGGCCACCTGGGATACACAGTTCTGTCGGGCTTTTCCGGAACACAGTTCC 1314	
DB	1213 GCCTTAGAGTTTCGTCTCTCAGGACTCATCAGTCTGAGCATTTGTCTAGTTTCAGGAGACT 1272	
QY	1315 GGATGTCAGCGCGAAGCGCATCAGCAACCCGCAACATACCGGCGACAGCCGGAACCTGCC 1374	
DB	1273 CTGGATCCAAGCCTTAGAGTTTCGTCTCACGGACTCATCAGTCTCTGAGCAATTGTCTAGT 1332	
QY	1375 GTGCCGCTGTGCAGATTAAATGACAGCGGTGGCGCTGGGATATTACGTCAGCGAGGAGC 1434	
DB	1333 TCAGGAGACTCTGGATCCAAGCCTTAGAGTTTCGTCTCACGGACTCATCAGTCTCTGAGC 1392	
QY	1435 GGTATCTCTGGCTGGATCGCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGCGGG- 1493	
DB	1393 ATTGTCTAGAGCGGCGCCACCGCGGTGGAGTCCAGCTTTTGTTCCTTTAGTGAGGCT 1452	
QY	1494 -----CGCGCTTGGCGTAATCATGTCATAGCTGTTTCTGTGTGAATTTGTTATCCGC 1547	
DB	1453 TAAATTCAGCTTGGCGTAATCATGTCATAGCTGTTTCTGTGTGAATTTGTTATCCGC 1512	
QY	1548 TCACAATTTCCACACAACATACGAGCGGAAGCATAAAGTGTAAAGCCTGGGGTGCTTAAT 1607	
DB	1513 TCACAATTTCCACACAACATACGAGCGGAAGCATAAAGTGTAAAGCCTGGGGTGCTTAAT 1572	
QY	1608 GAGTGAGCTAACTCACTAATTAATTGGGTGGGTCTCATCTGCCGCTTTTCAAGTCGGGAAC 1667	
DB	1573 GAGTGAGCTAACTCACTAATTAATTGGGTGGGTCTCATCTGCCGCTTTTCAAGTCGGGAAC 1632	
QY	1668 TGTCTGTCCAGCTGATTAATGMAATCGGCCAACCGCGGGGAGAGCGGTTTTCGCTATTG 1727	
DB	1633 TGTCTGTCCAGCTGATTAATGMAATCGGCCAACCGCGGGGAGAGCGGTTTTCGCTATTG 1692	
QY	1728 GCGCTCTTTCCGCTCTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 1787	
DB	1693 GCGCTCTTTCCGCTCTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 1752	
QY	1788 CGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATCCACAGAAATCAGGGATACGCGAG 1847	

1753	CGGTATCAGCTCACTCAAAGCGCGTAAATACGGTTATCCACAGAAATCAGGGGATAAACGCAG	1818
1848	GAAGAACAATGTGAGCAAAAAGGCCAGCAAAAGGCCAGGAACCGGTAAAAAAGGCCGCGTTGC	1907
1813	GAAGAACAATGTGAGCAAAAAGGCCAGCAAAAGGCCAGGAACCGGTAAAAAAGGCCGCGTTGC	1872
1908	TGGCGTTTTTCCATAGGCTCGCCCCCTGACGAGCATCAAAAATTCGAGCGTCAAGTC	1967
1873	TGGCGTTTTTCCATAGGCTCGCCCCCTGACGAGCATCAAAAATTCGAGCGTCAAGTC	1932
1968	AGAGTGGCGAAACCCGACAGGACTATAAAGATACCAAGGGGTTTCCCCCTGGAAAGCTCCC	2027
1933	AGAGTGGCGAAACCCGACAGGACTATAAAGATACCAAGGGGTTTCCCCCTGGAAAGCTCCC	1992
2028	TCGTGCGCTCTCCTGTTCCGACCCCTGCCGTTACCGGATACCTGTCCGCGCTTTCTCCCTT	2087
1993	TCGTGCGCTCTCCTGTTCCGACCCCTGCCGTTACCGGATACCTGTCCGCGCTTTCTCCCTT	2052
2088	CGGGAAGCTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGTCTCG	2147
2053	CGGGAAGCTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGTCTCG	2112
2148	TTGCTCCAAAGCTGGCGTGTGCAGAAACCCCGGTTTCAGCCCGACCGCTGGCGCCTTAT	2207
2113	TTGCTCCAAAGCTGGCGTGTGCAGAAACCCCGGTTTCAGCCCGACCGCTGGCGCCTTAT	2172
2208	CCGTTAACTATCGTCTTTGAGTCCAAACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAG	2267
2173	CCGTTAACTATCGTCTTTGAGTCCAAACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAG	2232
2268	CCACTGGTAAACAGATTTAGCAGAGCGAGGTATGTAGGGCGTGTACAGAGTTCTTTGAAGT	2327
2233	CCACTGGTAAACAGATTTAGCAGAGCGAGGTATGTAGGGCGTGTACAGAGTTCTTTGAAGT	2292
2328	GGTGGCCTAACTACGGCTACACTAGAGGACAGTATTTTGGTATCTCGCGCTCTGCTGAAGC	2387
2293	GGTGGCCTAACTACGGCTACACTAGAGGACAGTATTTTGGTATCTCGCGCTCTGCTGAAGC	2352
2388	CAGTTACCTTTCCGAAAAAGAGTTGGTAGTCTCTGATCCGGCAAAACAAACCAACCGCTGGTA	2447
2353	CAGTTACCTTTCCGAAAAAGAGTTGGTAGTCTCTGATCCGGCAAAACAAACCAACCGCTGGTA	2412
2448	CGCGTGGTTTTTTGTTTGAAGCAGCAGATTTACGCGCAGAAAAAAAGGATCTCAAGAAG	2507
2413	CGCGTGGTTTTTTGTTTGAAGCAGCAGATTTACGCGCAGAAAAAAAGGATCTCAAGAAG	2472
2508	ATCCTTTGATCTTTTCTACGGGCTTCACGCTCAGTGGAAAGAAACCTCAGTTTAAGGA	2567
2473	ATCCTTTGATCTTTTCTACGGGCTTCACGCTCAGTGGAAAGAAACCTCAGTTTAAGGA	2532
2568	TTTTGGTCAATGAGATTAATAAAGAGTCTTCACTAGATCTCTTTAAATTTAAAAATGAA	2627
2533	TTTTGGTCAATGAGATTAATAAAGAGTCTTCACTAGATCTCTTTAAATTTAAAAATGAA	2592
2628	GTTTTAAATCAAATCTAAAGTATATACAGTAAACTCGGTCTGACAGTTACCAATGCTTAA	2687
2593	GTTTTAAATCAAATCTAAAGTATATACAGTAAACTCGGTCTGACAGTTACCAATGCTTAA	2652
2688	TCAGTAGGCACCTATCTACGCGATCTGTCTATTTTGGTTCAATCATAGTTTGCCTGACTCC	2747
2653	TCAGTAGGCACCTATCTACGCGATCTGTCTATTTTGGTTCAATCATAGTTTGCCTGACTCC	2712
2748	CCGTGCTGTAGATAAATACGATACGGAGGGCTTACCATCTGGCCCCAGTGGCTGCAATGA	2807
2713	CCGTGCTGTAGATAAATACGATACGGAGGGCTTACCATCTGGCCCCAGTGGCTGCAATGA	2772
2808	TACCGCAGAGCCACGCTCAACCGGCTCCAGATTTTATCAGCAATAAACCAGCCAGCCGGAA	2867
2773	TACCGCAGAGCCACGCTCAACCGGCTCCAGATTTTATCAGCAATAAACCAGCCAGCCGGAA	2832
2868	GGGCCGAGCGCAGAGTGGTCTTGCAACTTTTATCCGCTCTCATCTCAGTCTATTAATTTGTT	2927
2833	GGGCCGAGCGCAGAGTGGTCTTGCAACTTTTATCCGCTCTCATCTCAGTCTATTAATTTGTT	2892

Qy	2928	GCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAAATAGTTTTGCGCAACGTTGTGGCATTG	2987
Db	2893	GCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAAATAGTTTGCGCAACGTTGTGGCATTG	2952
Qy	2988	CTACAGGCATCGTGGTGTCAACGCTCGTCTGTTGGTATGGCTTCATTACAGCTCCGGTTCCC	3047
Db	2953	CTACAGGCATCGTGGTGTCAACGCTCGTCTGTTGGTATGGCTTCATTACAGCTCCGGTTCCC	3012
Qy	3048	AACGATCAAGCGCAGATTACATGATCCCCCATGTTGTGCAAAAAGGGTTAGCTCCTTCG	3107
Db	3013	AACGATCAAGCGCAGATTACATGATCCCCCATGTTGTGCAAAAAGGGTTAGCTCCTTCG	3072
Qy	3108	GTCTCCGATCGTGTTCAGAATAAGTTGGCCGACGTGTTATCACTCATGGTTATGGCAG	3167
Db	3073	GTCTCCGATCGTGTTCAGAATAAGTTGGCCGACGTGTTATCACTCATGGTTATGGCAG	3132
Qy	3168	CACTGCATAACTCTCTTACTGTGCATGCCATCCGTAAAGATGCTTTTCTGTGACCTGGTGAGT	3227
Db	3133	CACTGCATAACTCTCTTACTGTGCATGCCATCCGTAAAGATGCTTTTCTGTGACCTGGTGAGT	3192
Qy	3228	ACTCAACCAGTCAATCTCTGAGAATAGTGTATGCGCGCACCGAGTTGCTCTTGGCCCGCGT	3287
Db	3193	ACTCAACCAGTCAATCTCTGAGAATAGTGTATGCGCGCACCGAGTTGCTCTTGGCCCGCGT	3252
Qy	3288	CAATACGGGANAATACCGGGCCAATAGCAGAACTTTTAAAAAGTGCTCATCTTGGAAAAAC	3347
Db	3253	CAATACGGGANAATACCGGGCCAATAGCAGAACTTTTAAAAAGTGCTCATCTTGGAAAAAC	3312
Qy	3348	GTCTCTCGGGCGAAACCTCTCAAGGATCTTACCGCTGTAGATCCAGTTCAGATGTAAAC	3407
Db	3313	GTCTCTCGGGCGAAACCTCTCAAGGATCTTACCGCTGTAGATCCAGTTCAGATGTAAAC	3372
Qy	3408	CCACTCGTCACCCAACTGATCTTACAGCATCTTTTACTTTTACACAGCGTTCTGGGGTGAG	3467
Db	3373	CCACTCGTCACCCAACTGATCTTACAGCATCTTTTACTTTTACACAGCGTTCTGGGGTGAG	3432
Qy	3468	CAAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACAACGAAATGTGAA	3527
Db	3433	CAAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACAACGAAATGTGAA	3492
Qy	3528	TACTCATCTCTCTCTTTTCAATATNTTGAAGCATTTATCAGGTTATTGTCTCATGA	3587
Db	3493	TACTCATCTCTCTCTTTTCAATATNTTGAAGCATTTATCAGGTTATTGTCTCATGA	3552
Qy	3588	GCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTTC	3647
Db	3553	GCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTTC	3612
Qy	3648	CCCGAAAAAGTGC 3659	
Db	3613	CCCGAAAAAGTGC 3624	
RESULT 5			
A20702			
LOCUS	A20702	3681 bp	DNA linear PAT 22-AUG-1994
DEFINITION	pKSEL5 DNA sequence.		
ACCESSION	A20702		
VERSION	A20702.1	GI:583404	
KEYWORDS	.		
SOURCE	synthetic construct		
ORGANISM	synthetic construct		
REFERENCE	other sequences; artificial sequences.		
AUTHORS	1 (bases 1 to 3681)		
TITLE	/organism="synthetic construct"		
JOURNAL	USE AS AN IMMUNOGEN AND VACCINE		
FEATURES	Patent: WO 9113155-A 6 05-SEP-1991;		
source	Location/Qualifiers		
	1..3681		
	/mol type="unassigned DNA"		

ORIGIN		/db_xref="taxon:32630"	
Query Match		72.1%; Score 2639; DB 6; Length 3681;	
Best Local Similarity		83.5%; Pred. No. 0;	
Matches 3075; Conservative		0; Mismatches 580; Indels 28; Gaps 6;	
QY	3	AAATTGTAAGCGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCAT	62
DB	1	AAATTGTAACGGTTAAATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCAT	60
QY	63	TTTTTAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGA	122
DB	61	TTTTTAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGA	120
QY	123	TAGGTTGAGTGTGTTCCAGTTTGGAAACAAAGATCCCACTATTAAGAACGTGGACTCCA	182
DB	121	TAGGTTGAGTGTGTTCCAGTTTGGAAACAAAGATCCCACTATTAAGAACGTGGACTCCA	180
QY	183	ACGTCAAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCT	242
DB	181	ACGTCAAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCT	240
QY	243	AATCAAGTTTTTGGGTCAGGTCGCGTAAAGCACTAAATCGGAACCTTAAAGGAGGCC	302
DB	241	AATCAAGTTTTTGGGTCAGGTCGCGTAAAGCACTAAATCGGAACCTTAAAGGAGGCC	300
QY	303	CCGGATTTAGACTTGAACGGGAAGACCGCGAAACGTGGCGAAGAGGAGGAAGAAG	362
DB	301	CCGGATTTAGACTTGAACGGGAAGACCGCGAAACGTGGCGAAGAGGAGGAAGAAG	360
QY	363	CGAAAGGAGCGGGCGTAGGGCGCTGGCAAGTGTAGCGGTACAGCTGGCGGTAAACACCA	422
DB	361	CGAAAGGAGCGGGCGTAGGGCGCTGGCAAGTGTAGCGGTACAGCTGGCGGTAAACACCA	420
QY	423	CACCGCGCGCTTAATGCGCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCA	482
DB	421	CACCGCGCGCTTAATGCGCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTACGCA	480
QY	483	ACTGTTGGAGGGCGATCGGTGGCGCTCTTCGCTATTAGCCAGCTGGCGGAAGGGG	542
DB	481	ACTGTTGGAGGGCGATCGGTGGCGCGCTCTTCGCTATTAGCCAGCTGGCGGAAGGGG	540
QY	543	GATGTGCTCAAGCGGATTAAGTTGGGTAAACCCAGGGTTTTCCAGTCAACAGCTGTGA	602
DB	541	GATGTGCTCAAGCGGATTAAGTTGGGTAAACCCAGGGTTTTCCAGTCAACAGCTGTGA	600
QY	603	AAACGACGCCAGTGAAGCGCGCTCGTTCAATTCACGTTTTTTGAAACCCGTGGAGGACGGG	662
DB	601	AAACGACGCCAGTGAATTTGTAATACGACTCACTATAGGGCGAATTTGGAGCTCCACCGCG	660
QY	663	AGACTCGCGTGCAATGTGTTTACAGGGTGATGGAGCAGATGAAGATGTCGACACGC	722
DB	661	GTGGCGCGCGCTCTAGTATGGTGCATCTCAGTACAATCTGCTCTGATGCGGCATAGTTA	720
QY	723	TGCAGAACACGCGCTAGATTAAACCTTAGAAGATAATCATATTGTGACGTACGTTAAAG	782
DB	721	AGCCAGTATATACATCCGCTATCGTACGTACGTGGGTCTGGCTGCCGCCGACACCC	780
QY	783	ATAATCATGCGTAAATTTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAAATG	842
DB	781	GCCAAACACCGCTGACGGCCCTGACGGGCTGTCTGCTCCGGCATCCGCTTACAGACA	840
QY	843	GGTACCGGGCCCCCTCGAGTTCAGCGTATCGA-----TAAGCTTGATATCAATTC	896
DB	841	AGCTGTGACCGCTCCGGGAGCTGATGTGTAGAGGTTTTTACCGTCACTACCGGAAACG	900
QY	897	CTGACAGCCGGGGATCCACTAGTTCTAGAGGGCGCGCCACCGCGTGGAGCTCCAGCTT	956
DB	901	CCGAGGACAGTAAGGTCCGATGCTTTGTGAGCAATTCGTCCCTTAAGTAAGCAATTGCTG	960
QY	957	TTGTTCCCTTTAGTGAGGGTTAATTAGATCCCAT-----GCGTCAATTTTACCGACAG	1009

DB	961	TAAAGTCGTCACTGTGCGGATCACCGTTCCAGTAGCGACAGAGCAATTTGATTGGTAAA	1020
QY	1010	TATCTTTCTAGGGTTAAATCTAGCTCATCAGGATCATATCGTCGGGTCTTTTTCGCGCT	1069
DB	1021	TTTCAGAGAAAGATCGGAGGAAGATCAATACATAAAGAGTTGAACCTTCITTTGTGTCT	1080
QY	1070	CAGTCATCCGCCCAAGCTGCGCTATCTGGGCATCTGGGAGGAAGAGCCGTCGCTTTTC	1129
DB	1081	TCGACATGGTAAATCTCATGTTTGAATGGCCCTAGAGGATCCCGCCCAAGCTTGCATGCC	1140
QY	1130	CGCGAGGTTGAACGGGCATGGAAGAGTTTTCGCCAGGATGACTCTGCTGATGATGACGT	1189
DB	1141	TGCAGTCCGACTCTAGAGGATCCCGCAGCTCGACGCCATTAATTAATTTTCCGTAAT	1200
QY	1190	TCAGGAAAACGACAGTTTACCATGATATTTCGGGAAGGTGTGGCATGCAAGCTTTAA	1249
DB	1201	TCAGGCGCTTCATGATGAGACAGGCGGTTT--GAATGTTGACGGGATGAACATAAATAG	1258
QY	1250	CGGTGAACCTGTTCTGTTTCAGGGCACCTGGGATACCAAGTTCTGTCGCGGCTTTTCCGACACA	1309
DB	1259	CAATGACGGCAGCAATAAATCAACAGG--AGCAGGAAGCGAGGGTATCCCAACAAGT	1315
QY	1310	GTTCGGATGTTGACGCCGAGCGCATCAGCAACCCGAAACAATACCGGGGACAG-----	1363
DB	1316	CCAGCGTACCATAAACGCAAGCCTCAACGCGACGACGAGACGAGAGCGGTCAGTAGCAA	1375
QY	1364	CGGAACCTGCGTGGGTGTCAGATTAAATGACAGCGGTGCGGCTGGGATATTACGT	1423
DB	1376	TCCAAACTTTGTTACTGTCAGAAATCGAAATCATCTTCGGTTAAATCCAAACGGCAG	1435
QY	1424	CAGCAGGACGGGTAT----CCTGGCTGGATGCCGCAAGAAATGGACATGCCCTTACCCGTTG	1479
DB	1436	AAGCTGAATGAGAATTCGACCTCGAGGGGGCGCCGTCACCGACTTTTGTTCCTTTA	1495
QY	1480	AGTTACCGCGGGCGCGCTTGGCGTAAATCATGTCATAGCTGTTTCTGTGTGAAATTG	1539
DB	1496	GTGAGGGTTAAATTCGAGCTTGGCGTAAATCATGTTTCATAGCTGTTTCTGTGTGAAATTG	1555
QY	1540	TTATCCGCTCAAAATTCACACACATACGAGCGGAAGCATAAAGTGTAAAGCTGGGG	1599
DB	1556	TTATCCGCTCAAAATTCACACACATAGGAGCGGAAGCATAAAGTGTAAAGCTGGGG	1615
QY	1600	TGCTTAATGAGTGAGCTAACTCACAATTAATTCGTTGCGTCACTGCCCGCTTTCCAGTTC	1659
DB	1616	TGCTTAATGAGTGAGGTAACTCACAATTAATTCGTTGCGTCACTGCCCGCTTTCCAGTTC	1675
QY	1660	GGGAAACCTGTGTCGACGTGCATTAATGAATCGGCCAAACGCGCGGGAGGCGGTTT	1719
DB	1676	GGGAAACCTGTGTCGACGTGCATTAATGAATCGGCCAAACGCGCGGGAGGCGGTTT	1735
QY	1720	CGGTATTCGGCGCTTTCGCTTCTCGCTCACTGACCTCGCTCGCTCGCTCGCTCGCT	1779
DB	1736	CGGTATTCGGCGCTTTCGCTTCTCGCTCACTGACCTCGCTCGCTCGCTCGCTCGCT	1795
QY	1780	CGCGCAGCGGTATCAGCTCACTCAAAGCGGTAAATCGGTTATCCACAGAATCAGGGA	1839
DB	1796	CGCGCAGCGGTATCAGCTCACTCAAAGCGGTAAATCGGTTATCCACAGAATCAGGGA	1855
QY	1840	TAAACGAGAAAGAACATGTGAGCAAAAGCGCAAAAGCGCAGGAACCGTAAAGAGC	1899
DB	1856	TAAACGAGAAAGAACATGTGAGCAAAAGCGCAGCAAAAGCGCAGGAACCGTAAAGAGC	1915
QY	1900	CGCGTTCGCGGTTTTTCCATAGCTCGCCCTCGACGAGCATCAAAATAATCGACG	1959
DB	1916	CGCGTTCGCGGTTTTTCCATAGCTCGCCCTCGACGAGCATCAAAATAATCGACG	1975
QY	1960	CTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACCAAGCGTTTCCCGCTGG	2019
DB	1976	CTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACCAAGCGTTTCCCGCTGG	2035
QY	2020	AAGCTCCCTCGTGGCTTCTGTCGACCTGCGCTTACCGATACCTGTCGCGCTT	2079
DB	2036	AAGCTCCCTCGTGGCTTCTGTCGACCTGCGCTTACCGATACCTGTCGCGCTT	2095

Qy	2080	TCTCCCTTCGGGAAGCGTGGCGCTTTTCTCATAGCTCA CGCTGTAGGTATCTCAAGTTCCGT	2139
Db	2096	TCTCCCTTCGGGAAGCGTGGCGCTTTTCTCAATGCTCA CGCTGTAGGTATCTCAAGTTCCGT	2155
Qy	2140	GTAGGTGGTTGGCTTCCAAGCTGGCGCTGTGTGCAGAACCCCGTTGAGCCCGCCGACCGCTG	2199
Db	2156	GTAGGTGGTTGGCTTCCAAGCTGGCGCTGTGTGCAGAACCCCGTTGAGCCCGCCGACCGCTG	2215
Qy	2200	CGCCTTATCCGGTAACCTATCGTCTTGAGTCCAAACCCGGTAAGACACAGACTTATTCGCCACT	2259
Db	2216	CGCCTTATCCGGTAACCTATCGTCTTGAGTCCAAACCCGGTAAGACACAGACTTATTCGCCACT	2275
Qy	2260	GGCAGCAGCCACTGGTAAACAGGATTAGCAGACGAGGTATGTAGCGCGTGCTACAGAGTT	2319
Db	2276	GGCAGCAGCCACTGGTAAACAGGATTAGCAGACGAGGTATGTAGCGCGTGCTACAGAGTT	2335
Qy	2320	CTTGAAAGTGGTGGCTTAACCTACGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCT	2379
Db	2336	CTTGAAAGTGGTGGCTTAACCTACGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCT	2395
Qy	2380	GCTCAAGCAGTTACCTTCGGAAAAAGTTGGTAGCTCTTGATCCGGCAAAACAACAC	2439
Db	2396	GCTCAAGCAGTTACCTTCGGAAAAAGTTGGTAGCTCTTGATCCGGCAAAACAACAC	2455
Qy	2440	CGCTGGTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGATTACGCGCAGAAAAAAGGATC	2499
Db	2456	CGCTGGTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGATTACGCGCAGAAAAAAGGATC	2515
Qy	2500	TCAAGAAAGATCCTTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAATCAACG	2559
Db	2516	TCAAGAAAGATCCTTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAATCAACG	2575
Qy	2560	TTAAGGATTTTGGTCATGAGATTATCAAAAAGATCTTCACCTAGATCCCTTTTAAATTA	2619
Db	2576	TTAAGGATTTTGGTCATGAGATTATCAAAAAGATCTTCACCTAGATCCCTTTTAAATTA	2635
Qy	2620	AAAAATGAAGTTTTAAATCAATCTAAAGTATATATAGTAAACTTTGGTCTGCACAGTTACCA	2679
Db	2636	AAAAATGAAGTTTTAAATCAATCTAAAGTATATATAGTAAACTTTGGTCTGCACAGTTACCA	2695
Qy	2680	ATGCTTAAATCAGTGAGGCACTATCTCAGCGATCTGCTATTTGGTTTCATTCATAGTTGC	2739
Db	2696	ATGCTTAAATCAGTGAGGCACTATCTCAGCGATCTGCTATTTGGTTTCATTCATAGTTGC	2755
Qy	2740	CTGACTCCCGTGGTGTAGATACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGC	2799
Db	2756	CTGACTCCCGTGGTGTAGATACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGC	2815
Qy	2800	TGCAATGATACCGCGAGACCCAGCCTCAACGGGTCCAGATTATATACGAATAAACACAGCC	2859
Db	2816	TGCAATGATACCGCGAGACCCAGCCTCAACGGGTCCAGATTATATACGAATAAACACAGCC	2875
Qy	2860	AGCCGGAAGGCGGAGCGAGAGTGTCTCTGCAACTTTATTCGGCTCCATCCAGTCTAT	2919
Db	2876	AGCCGGAAGGCGGAGCGAGAGTGTCTCTGCAACTTTATTCGGCTCCATCCAGTCTAT	2935
Qy	2920	TAAATGTTGCCGGAAAGTGAAGTATGTTCCGCAGTTAAATAGTTTGGCGCAACGGTTGT	2979
Db	2936	TAAATGTTGCCGGAAAGTGAAGTATGTTCCGCAGTTAAATAGTTTGGCGCAACGGTTGT	2995
Qy	2980	TGCATTTGCTACAGGCATCGTGGTGTCAACGCTCGTCTGTTGGTATGGCTTCATTCAGCTC	3039
Db	2996	TGCATTTGCTACAGGCATCGTGGTGTCAACGCTCGTCTGTTGGTATGGCTTCATTCAGCTC	3055
Qy	3040	CGGTTCCCAACGATCAAGGGGAGTTACATGATCCCCCATGTGTGTGCAAAAAAGCGGTTAG	3099
Db	3056	CGGTTCCCAACGATCAAGGGGAGTTACATGATCCCCCATGTGTGTGCAAAAAAGCGGTTAG	3115
Qy	3100	CTCCTTCGGTCTCCGATCGTTGTGAGAAAGTAAAGTTGGCGCGAGTTATCATCTATGGT	3159
Db	3116	CTCCTTCGGTCTCCGATCGTTGTGAGAAAGTAAAGTTGGCGCGAGTTATCATCTATGGT	3175

Qy	3160	TATGGCAGCACTGCATTAATTTCTCTTACTGTGCATGCCATCCGTGAAGATGCTTTTCTGTGCAC	3212
Db	3176	TATGGCAGCACTGCATTAATTTCTCTTACTGTGCATGCCATCCGTGAAGATGCTTTTCTGTGCAC	3235
Qy	3220	TGGTGAGTACTCAACCAAGTCATTCTTGAGAAATAGTGTATGTCGGCGACCGAGTTGCTCTTG	3279
Db	3236	TGGTGAGTACTCAACCAAGTCATTCTTGAGAAATAGTGTATGTCGGCGACCGAGTTGCTCTTG	3295
Qy	3280	CCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCAT	3339
Db	3296	CCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCAT	3355
Qy	3340	TGGAATAACGTTCTTCGGGGCGAAAACTCTCAAGATCTTACCGCTGTGTAGATCCAGTTC	3399
Db	3356	TGGAATAACGTTCTTCGGGGCGAAAACTCTCAAGATCTTACCGCTGTGTAGATCCAGTTC	3415
Qy	3400	GATGTAAACCCACATCGTGCACCCCACTCATCTTCAGAGATCTTTTACTTTTACCAGCGTTTC	3459
Db	3416	GATGTAAACCCACATCGTGCACCCCACTCATCTTCAGAGATCTTTTACTTTTACCAGCGTTTC	3475
Qy	3460	TGGGTGAGCAAAAACAGGAAGGCAAAATGCGCCAAAAAGGGAATAAGGCGCAGACGGAA	3519
Db	3476	TGGGTGAGCAAAAACAGGAAGGCAAAATGCGCCAAAAAGGGAATAAGGCGCAGACGGAA	3535
Qy	3520	ATGTTGAATACTCATACTCTCTTTTCAATATATTTGAAGCATTTATCAGGGTTATTG	3579
Db	3536	ATGTTGAATACTCATACTCTCTTTTCAATATATTTGAAGCATTTATCAGGGTTATTG	3595
Qy	3580	TCTCATGAGCGGATACATATTTGAAATGATTTTGAAGAAATAAACAATAGGGGTTCCGCG	3639
Db	3596	TCTCATGAGCGGATACATATTTGAAATGATTTTGAAGAAATAAACAATAGGGGTTCCGCG	3655
Qy	3640	CACATTTCCCGAAAAAGTGCAC 3662	
Db	3656	CACATTTCCCGAAAAAGTGCAC 3678	
RESULT 6			
I15651			
LOCUS	I15651	3681 bp	DNA linear PAT 02-APR-1996
DEFINITION	Sequence 6 from patent US 5470573.		
ACCESSION	I15651		
VERSION	I15651.1		
KEYWORDS	GI:1250559		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 3681)		
AUTHORS	Lubitz, W. and Szostak, M.P.		
TITLE	Immunogens comprising the non-lytic membrane spanning domain of bacteriophages MS2 or PhiX174		
JOURNAL	Patent: US 5470573-A 6 28-NOV-1995;		
FEATURES	Location/Qualifiers		
source	1. 3681		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN			
Query Match	72.1%;	Score 2639;	DB 6; Length 3681;
Best Local Similarity	83.5%;	Pred. No. 0;	
Matches 3075;	Conservative 0;	Mismatches 580;	Indels 28; Gaps 6;
Qy	3	AAATTGTACGGTTAATATTTTGTAAATTCGGGTTAAATTTTGTAAATCAGTTCAT	62
Db	1	AAATTGTAAACGGTTAATATTTTGTAAATTCGGGTTAAATTTTGTAAATCAGTTCAT	60
Qy	63	TTTTTAACCAATAGGCGGAAATCGGCAAAATCCCTTATAATCAAAGAAATAGACCGAGA	122
Db	61	TTTTTAACCAATAGGCGGAAATCGGCAAAATCCCTTATAATCAAAGAAATAGACCGAGA	120
Qy	123	TAGGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAAGTGGACTCCA	182
Db	121	TAGGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAAGTGGACTCCA	180

183 QY ACCTCAAAGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACTGTGAACCATCACCCCT 242
184 DB |||||
181 QY AGCTCAAAGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACTGTGAACCATCACCCCT 240
182 DB |||||
243 QY AATCAAGTTTTTTGGGGTTCAGGTCGCGTAAAGCACTAAATCGGAACCCCTAAAGGGAGCC 302
244 DB |||||
241 QY AATCAAGTTTTTTGGGGTTCAGGTCGCGTAAAGCACTAAATCGGAACCCCTAAAGGGAGCC 300
242 DB |||||
303 QY CCGGATTTAGAGCTTGACGGGAAAGCCGGCGAAAGCTGCGGAAAGGAGGAGGAGAAAG 362
304 DB |||||
301 QY CCGGATTTAGAGCTTGACGGGAAAGCCGGCGAAAGCTGCGGAAAGGAGGAGGAGAAAG 360
302 DB |||||
363 QY CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACTGCGTAAACCA 422
364 DB |||||
361 QY CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACTGCGTAAACCA 420
362 DB |||||
423 QY CACCGCGCGCTTAAATGCGCGCTACAGGGCGCGTCCCAATTCGCGATTCAGGTCGCGCA 482
424 DB |||||
421 QY CACCGCGCGCTTAAATGCGCGCTACAGGGCGCGTCCCAATTCGCGATTCAGGTCGCGCA 480
422 DB |||||
483 QY ACTGTTGGGAAGGCGATCGGTCGGGCTCTTCGCTATTACGCCAGCTGGCGAAGGGG 542
484 DB |||||
481 QY ACTGTTGGGAAGGCGATCGGTCGGGCTCTTCGCTATTACGCCAGCTGGCGAAGGGG 540
482 DB |||||
543 QY GATGTGCTGCAAGGCGATTAAGTTGGGTAAACCGCAGGGTTTTCCAGTCAACGACGTTGTA 602
544 DB |||||
541 QY GATGTGCTGCAAGGCGATTAAGTTGGGTAAACCGCAGGGTTTTCCAGTCAACGACGTTGTA 600
542 DB |||||
603 QY AAACGAGCGCGATGAGCGCGCTCGTTCAATTCAGCTTTTTTGAACCCCGTGGAGGACGGG 662
604 DB |||||
601 QY AAACGAGCGCGATGAGCTGAATTAACCGTCAAGAAAGATAATCATATTGTGACGTACGTTAAAG 782
602 DB |||||
721 QY AGCCAGTATATACACTCCGCTATCGTACGTGACGTGGGTTCATGGCTGCGCCCGACACCC 780
722 DB |||||
783 QY ATAATCATGCGTAAATTTGACGATGCGATCTGTAAATACGATCACTATAGGGCGAAATTG 842
784 DB |||||
781 QY GCCAACACCGCTGACGGCGCTGACGGGCTTGTCTGCTCCGGCATCCGCTTACAGACA 840
782 DB |||||
843 QY GGTACGGGCGCCCTCGAGGTCGACGATCGA- ----TAAGCTTGATATCGAATTC 896
844 DB |||||
841 QY AGCTGTGACCGTCTCCGGAGCTGCATGTGTGAGAGGTTTTCCCGTCAATCACCGAAAGC 900
842 DB |||||
897 QY CTGACGCGCGGGGATCCACTAGTTCTAGAGCGGCGCCACCGCGGTGGAGTCCAGGTT 956
898 DB |||||
901 QY CCGGAGGAGTAAGGTCGATGCTTTGTGAGCAATTCGTCCTTAAAGTAAGCAATTGCTG 960
902 DB |||||
957 QY TTGTTCCCTTTAGTGAGGGTTAAATTAGATCCAT- ----GCGTCAATTTTACGAGAC 1009
958 DB |||||
961 QY TAAAGTCGTCACTGTGCGGATCAGCGCTTCCAGTAGCGACAGAAAGCAATTGATTTGTTAA 1020
962 DB |||||
1010 QY TATCTTTCTAGGGTTAAATCTAGCTGCATCAGGATCATATCTGTCGGGTCTTTTTCGGCT 1069
1011 DB |||||
1021 QY TTTGAGAGAAAGATCGCGAGGAAGATCAATACATAAAGAGTTGAACCTTCTTTGTTGTCT 1080
1022 DB |||||
1070 QY CAGTCATCGCCCAAGCTGCGCTATCTGGGCATCGGGAGGAGAAAGCCCGTGCCTTTTC 1129
1071 DB |||||
1081 QY TCGACATGGGTAATCTCATGTTTGAATGGCCCTAGAGGATCCGGCCAAGCTTGATGCC 1140
1082 DB |||||
1130 QY CCGGAGGTTGAAGCGGATGAAAGAGTTTCCGAGGATGATCTGCTGTGATTTGACGT 1189
1131 DB |||||
1141 QY TCGAGGTGACCTTAGAGGATCCCGGACGCTCGAGCGCAATTAATAATGTTTTTCCGTAAT 1200
1142 DB |||||
1190 QY TGAGCGMAAACGACGCTTTACCATGATTCGGGAAGGTGTGGCAATGCGACGCTTTAA 1249
1191 DB |||||
1201 QY TCAGCGCTTTCATGATGAGACAGGCGGCTT- --GAATGTTGAGGAGTAACTAATTAAG 1258
1202 DB |||||

1250 QY CGGTGAACCTGTTCTGTTAGGCGCACCTGGGATACCAAGTTCTGTCGCGGCTTTTCCGACACA 1309
1251 DB |||||
1259 QY CAATCAGCGCAGCAATAAATCAACAGG- --AGCAGGAAAGCGAGGGTATCCCAAAAGT 1315
1260 DB |||||
1310 QY GTTCCGGAATGTCAGCCCGGAAGCGCATCAGCAACCCGAAACAAATACCGGGGACAG- ---- 1363
1311 DB |||||
1316 QY CCAGCGTACCATAAAGCAAGCCTCAACGCGACGACGAGCAGCAGAGCGGTGAGTAGCAA 1375
1317 DB |||||
1364 QY CCGGAACCTGCGTCCGGTGTGAGATTAAATGACAGCGGTGCGGCTGGGATATTACGT 1423
1365 DB |||||
1376 QY TCCAAACTTTGTTACTCGTCAGAAATCGAAATCATCTTCGTTAAATCCAAAAACGGCAG 1435
1377 DB |||||
1424 QY CAGCAGGACCGGGTAT- ----CCTGCTGGATGCCGACAGAAATGGACATGGAATACCCCGTG 1479
1425 DB |||||
1436 QY AAGCCTGAATGAGAAATTCGACCTCGAGGGGGCGCGGTACCCAGCTTTTGTTCCTTTA 1495
1437 DB |||||
1480 QY AGTTACCGCGGGGCGCTTGGCGTAAATCATGTGTCATAGCTTTCTCTGTGTGAATTG 1539
1481 DB |||||
1496 QY GTGAGGGTTAAATTCGAGCTTTGGCGTAAATCATGTGTCATAGCTTTCTCTGTGTGAATTG 1555
1497 DB |||||
1540 QY TTATCCGCTCAAAATTCACACAACATACGAGCCGGAAGCATAAAAGTGTAAAGCCTGGGG 1599
1541 DB |||||
1556 QY TTATCCGCTCAAAATTCACACAACATAGGAGCCGGAAGCATAAAAGTGTAAAGCCTGGGG 1615
1557 DB |||||
1600 QY TGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTC 1659
1601 DB |||||
1616 QY TGCCTAATGAGTGAGTTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTC 1675
1617 DB |||||
1660 QY GGGAAACCTGTCGTCAGCTGCATTAAATGAATCGGCCAAACGCGCGGGGAGGCGGTTTT 1719
1661 DB |||||
1676 QY GGGAAACCTGTCGTCAGCTGCATTAAATGAATCGGCCAAACGCGCGGGGAGGCGGTTTT 1735
1677 DB |||||
1720 QY GCGTATTCGGGCGCTCTTCGCTCTCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1779
1721 DB |||||
1736 QY GCGTATTCGGGCGCTCTTCGCTCTCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1795
1737 DB |||||
1780 QY GCGGAGCGGTATCAGCTCACTCAAAGCGGTAAATCGGTTATCCACAGAAATCAGGGA 1839
1781 DB |||||
1796 QY GCGGAGCGGTATCAGCTCACTCAAAGCGGTAAATCGGTTATCCACAGAAATCAGGGA 1855
1797 DB |||||
1840 QY TAAACGAGAAAGACATGTGAGCAAAAGGCGAGCAAAAGGCGAGGACCGTAAAGAGC 1899
1841 DB |||||
1856 QY TAAACGAGAAAGAACATGTGAGCAAAAGGCGAGCAAAAGGCGAGGACCGTAAAGAGC 1915
1857 DB |||||
1900 QY CGCGTTGCTGCGGTTTTTCCATAGGCTCCGCGCCCTGACGAGCATCAAAAAATCGACG 1959
1901 DB |||||
1916 QY CGCGTTGCTGCGGTTTTTCCATAGGCTCGGCGCCCTGACGAGCATCAAAAAATCGACG 1975
1917 DB |||||
1960 QY CTCAGTCAAGGTGGCGAAACCCGACAGGACTATAAAGATACAGGGGTTTTCCCGCTGG 2019
1961 DB |||||
1976 QY CTCAGTCAAGGTGGCGAAACCCGACAGGACTATAAAGATACAGGGGTTTTCCCGCTGG 2035
1977 DB |||||
2020 QY AAGTCCCTGCTGCGCTCTCTGTTCCGACCTGCGCTTACCGATACCTGTCGCGCTT 2079
2021 DB |||||
2036 QY AAGTCCCTGCTGCGCTCTCTGTTCCGACCTGCGCTTACCGATACCTGTCGCGCTT 2095
2037 DB |||||
2080 QY TCTCCCTTCGGGAAGCGGTGGCGCTTTCTCATAGCTCAAGCTGTAGGTATCTCAGTTCCGT 2139
2081 DB |||||
2096 QY TCTCCCTTCGGGAAGCGGTGGCGCTTTCTCAATGCTCAAGCTGTAGGTATCTCAGTTCCGT 2155
2097 DB |||||
2140 QY GTAGTCGTTTCGCTCAAAGCTGGGCTGTGTGACAGAACCCCGCTTCAGCCCGACCGCTG 2199
2141 DB |||||
2156 QY GTAGTCGTTTCGCTCAAAGCTGGGCTGTGTGACAGAACCCCGCTTCAGCCCGACCGCTG 2215
2157 DB |||||
2200 QY GCGCTTATCCGGTAACTATCGTCTGAGTCCAAACCCGCTAAGACACGACTTATCCCACT 2259
2201 DB |||||
2216 QY GCGCTTATCCGGTAACTATCGTCTGAGTCCAAACCCGCTAAGACACGACTTATCCCACT 2275
2217 DB |||||
2260 QY GCGACGCGCACTGGTAAACAGGATTAGCAGACGAGGTATGTAGCGGTGCTGTACAGAGTT 2319
2261 DB |||||
2276 QY GCGACGCGCACTGGTAAACAGGATTAGCAGACGAGGTATGTAGCGGTGCTGTACAGAGTT 2335
2277 DB |||||
2320 QY CTTGAAGTGGTGGCGCTAACTAGCGGCTACACTAGAAGGACAGTATTTGGTATCTCGGCTCT 2379
2321 DB |||||

```
Db 2336 CTTGAAGTGGTGGCCCTAACTACCGCTACACTAGAAGCAGTAGTATTGGTATCTGCGCTCT 2395
QY 2380 GCTGAAGCCAGTTACCTTCGGAAGAAAGAGTGGTAGCTCTTGATCCGGCAAAACAAACCCAC 2439
Db 2396 GCTGAAGCCAGTTACCTTCGGAAGAAAGAGTGGTAGCTCTTGATCCGGCAAAACAAACCCAC 2455
QY 2440 CGCTGTAGCGGTGGTTTTTTTGTGCAAGCAGCAGATPACCGCAGAGAAAAAAGGATC 2499
Db 2456 CGCTGTAGCGGTGGTTTTTTTGTGCAAGCAGCAGATPACCGCAGAGAAAAAAGGATC 2515
QY 2500 TCAAGAGAGTCTTTTCATCTTTTACGGGTCTGACGCTCAGTGGAACGAAACCTCAG 2559
Db 2516 TCAAGAGAGTCTTTTCATCTTTTACGGGTCTGACGCTCAGTGGAACGAAACCTCAG 2575
QY 2560 TTAAGGGATTTTGGTCATGAGATTATCAAAAGGATCTTACCTAGATCTTTTAAATTA 2619
Db 2576 TTAAGGGATTTTGGTCATGAGATTATCAAAAGGATCTTACCTAGATCTTTTAAATTA 2635
QY 2620 AAAATGAAGTTTAAATCAATCTAAAGTATATATAGTAAACTTGTCTGACAGATTACCA 2679
Db 2636 AAAATGAAGTTTAAATCAATCTAAAGTATATATAGTAAACTTGTCTGACAGATTACCA 2695
QY 2680 ATGCTTAATCAGTGAGGCCATCTCAGCGATCTCTATTTCTGTTCCATCATAGTTGC 2739
Db 2696 ATGCTTAATCAGTGAGGCCATCTCAGCGATCTCTATTTCTGTTCCATCATAGTTGC 2755
QY 2740 CTGACTCCCGCTCGTGTAGATACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGC 2799
Db 2756 CTGACTCCCGCTCGTGTAGATACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGC 2815
QY 2800 TGCATGATACCGCGAGACCCAGCTCACCAGTCTCAGATTTATCAGCAATAAACAGCC 2859
Db 2816 TGCATGATACCGCGAGACCCAGCTCACCAGTCTCAGATTTATCAGCAATAAACAGCC 2875
QY 2860 AGCCGAAGGGCGGACGCGAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTAT 2919
Db 2876 AGCCGAAGGGCGGACGCGAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTAT 2935
QY 2920 TAAATCTCCGGAAGCTAGAGTAGTAGTTCGCCAGTTAATAGTTTGCACAACTGT 2979
Db 2936 TAAATCTCCGGAAGCTAGAGTAGTAGTTCGCCAGTTAATAGTTTGCACAACTGT 2995
QY 2980 TGCATTTGTACAGGCATCGTGTGTACGCTCGTCTGTTGGTATGGCTTCAATCAGCTC 3039
Db 2996 TGCATTTGTACAGGCATCGTGTGTACGCTCGTCTGTTGGTATGGCTTCAATCAGCTC 3055
QY 3040 CGGTTCCCAACGATCAAGCGAGTTACATGATCCCTCCATGTTGCAAAAAGCGGTAG 3099
Db 3056 CGGTTCCCAACGATCAAGCGAGTTACATGATCCCTCCATGTTGCAAAAAGCGGTAG 3115
QY 3100 CTCCTTCGGTCTCCGATCGTGTGAGAGTAGTTCGCCAGTGTATCACTCATGT 3159
Db 3116 CTCCTTCGGTCTCCGATCGTGTGAGAGTAGTTCGCCAGTGTATCACTCATGT 3175
QY 3160 TATGGCAGCATGTCATAAATCTCTTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGAC 3219
Db 3176 TATGGCAGCATGTCATAAATCTCTTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGAC 3235
QY 3220 TGTGAGTACTCAACCAAGTCAATCTTGAGATAGTGTATCGGCGACCGAGTGTCTTTG 3279
Db 3236 TGTGAGTACTCAACCAAGTCAATCTTGAGATAGTGTATCGGCGACCGAGTGTCTTTG 3295
QY 3280 CCGCGGTCAATACGGGATAATACCGGCACATAGCAGAACTTTTAAAGTGTCTCATCAT 3339
Db 3296 CCGCGGTCAATACGGGATAATACCGGCACATAGCAGAACTTTTAAAGTGTCTCATCAT 3355
QY 3340 TGGAAACCGTCTTCGGGGCGAAAACTCTCAAGGATCTTACCCTGTGTGAGATCCAGTTC 3399
Db 3356 TGGAAACCGTCTTCGGGGCGAAAACTCTCAAGGATCTTACCCTGTGTGAGATCCAGTTC 3415
QY 3400 GATGTAAACCATCTGTGCAACCCAACTGATCTTTCAGCATCTTTTACCTTTTCCAGCGTTTC 3459
```

```
Db 3416 GATGTAAACCCACTCTGTGCAACCCAACTGATCTTTCAGCATCTTTTACCTTTCCACGCGTTTC 3475
QY 3460 TGGGTGAGCAAAAAACAGGAAGCAAAATGCCGCAAAAAAGGAATTAAGGGCGCACACGGAA 3519
Db 3476 TGGGTGAGCAAAAAACAGGAAGCAAAATGCCGCAAAAAAGGAATTAAGGGCGCACACGGAA 3535
QY 3520 ATGTTGAATACTCATCTCTCTTTTCAATATATTTGAAGCATTTATCAGGGTTATTG 3579
Db 3536 ATGTTGAATACTCATCTCTCTTTTCAATATATTTGAAGCATTTATCAGGGTTATTG 3595
QY 3580 TCTCATGACGGATACATATTTGAATGATTTAGAAAAATAAACAATAGGGTTCCGCG 3639
Db 3596 TCTCATGACGGATACATATTTGAATGATTTAGAAAAATAAACAATAGGGTTCCGCG 3655
QY 3640 CACATTTCCCGAAAGTGCAC 3662
Db 3656 CACATTTCCCGAAAGTGCAC 3678

RESULT 7
LOCUS A20700 5314 bp DNA linear PAT 22-AUG-1994
DEFINITION DMV1 DNA sequence.
ACCESSION A20700
VERSION A20700.1 GI:583402
KEYWORDS synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 5314)
AUTHORS other sequences; artificial sequences.
TITLE CARRIER-BOUND RECOMBINANT PROTEIN, PROCESS FOR PRODUCING IT AND ITS
JOURNAL USE AS AN IMMUNOGEN AND VACCINE
FEATURES Patent: WO 91/3155-A 4 05-SEP-1991;
LOCATION/Qualifiers
source
1..5314
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 71.6%; Score 2623.6; DB 6; Length 5314;
Best Local Similarity 83.4%; Pred. No. 0;
Matches 3059; Conservative 0; Mismatches 579; Indels 28; Gaps 6;
QY 20 ATTTTGTAAAATTCGCGTTAAATTTTGTAAATCAGCTCAATTTTAAACCAATAGGCC 79
Db 1651 ATTTTGTAAAATTCGCGTTAAATTTTGTAAATCAGCTCAATTTTAAACCAATAGGCC 1710
QY 80 GAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGATAGGGTTGAGTTGTT 139
Db 1711 GAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGATAGGGTTGAGTTGTT 1770
QY 140 CCAGTTTGAACAAGAGTCCACTATTAAGAACGTTGACTCCCAACGTCAGAGGCGCAAAA 199
Db 1771 CCAGTTTGAACAAGAGTCCACTATTAAGAACGTTGACTCCCAACGTCAGAGGCGCAAAA 1830
QY 200 ACCGTCATACGCGGATGCGCCACTACCTGTAACCAATCACTCAAGTTTTCGGGG 259
Db 1831 ACCGTCATACGCGGATGCGCCACTACCTGTAACCAATCACTCAAGTTTTCGGGG 1890
QY 260 TCGAGTGTCCGTAAGACACTAAATCGGAACCCCTAAAGGGAGGCCCCCGATTTAGAGCTTGA 319
Db 1891 TCGAGTGTCCGTAAGACACTAAATCGGAACCCCTAAAGGGAGGCCCCCGATTTAGAGCTTGA 1950
QY 320 CCGGGAAACCCCGCGCAACGTCGCGAGAAAGGAAGGAAAGGAAAGGAGAGCGGCGCT 379
Db 1951 CCGGGAAACCCCGCGCAACGTCGCGAGAAAGGAAGGAAAGGAGAGCGGCGCT 2010
QY 380 AGGGCGCTGCGCAAGTGTAGCGGTACGCTGCGGTAAACCAACACCCCGCGCTTAAT 439
Db 2011 AGGGCGCTGCGCAAGTGTAGCGGTACGCTGCGGTAAACCAACACCCCGCGCTTAAT 2070
```

QY 440 GCGCGCTACAGGCGCGTCCATTGCGCAATTCAGGCTCGCAACTGTTGGGAAGGCGA 499
Db |||||
2071 GCGCGCTACAGGCGCGTCCCATTCGCCATTACAGGCTAGCAACTGTTGGGAAGGCGA 2130
QY 500 TCGGTGCGGCGCTTTCGCTATTACGCCAGCTGGCGAAGAGGGGATGCTCTCAAGGCGA 559
Db |||||
2131 TCGGTGCGGCGCTTTCGCTATTACGCCAGCTGGCGAAGGGGATGCTCTCAAGGCGA 2190
QY 560 TTAAGTTGGTAAAGCGCAGGGTTTTCCAGTCACGACGTTGTAAGAACGACGCGCAGTGAG 619
Db |||||
2191 TTAAGTTGGTAAAGCGCAGGGTTTTCCAGTCACGACGTTGTAAGAACGACGCGCAGTGAA 2250
QY 620 CCGCGCTCGTTCATTTCACGCTTTTGAACCCGTGGAGACCGGCAGACTCGCGTGCAAT 679
Db |||||
2251 TTGTAATACGACTCACTATAGGCGGAATTGGAGCTCCACCGCGTGGCGCGCTCTAGT 2310
QY 680 GTGTTTACAGCGTGATGAGAGAGATGAAGATGCTCGACACGCTGAGCAACACGCGACTA 739
Db |||||
2311 ATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGATAGTTAAGCCAGTATATACACTC 2370
QY 740 GATTAAACCTAGAAAGATAATCATATTGAGCTAGTTAAAGATAATCATCGTAAAT 799
Db |||||
2371 CGCTATCGTACGTAAGCTGGGCTGAGGCTGCGGCCGACACCCGCCAACCACCGCTGAG 2430
QY 800 TGACGCTAGGATCTGTATACGACTCACTATAGGCGGAATTTGGGTACCGGGCCCCCCT 859
Db |||||
2431 CGCCCTGACGGCTTGTCTGCTCCCGCATCCGTTACAGACAAGCTGTGACCGTCTCG 2490
QY 860 CGAGTTCGACGATATCGA-----TAAGCTTGATATCGAATTCCTGACGCCCGGGGATC 913
Db |||||
2491 GGAGCTGCAATGCTCAGAGGTTTTACCGCTCATCACCGAAGCGCGAGGAGTAAAGGTC 2550
QY 914 CACTAGTTCTAGAGCGGCCACCGCGGTGAGCTCCAGCTTTGTTCCCTTTAGTGAG 973
Db |||||
2551 GGATGCTTTGTGAGCAATTCGCTCCCTTAAAGTAAGCAATTGCTGTAAGCTCACTGTC 2610
QY 974 GGTAAATAGATCCCAT-----CGGTCAATTTTACGCAGACTATCTTTCTAGGGTTAA 1026
Db |||||
2611 GGATCAGCGCTTCCAGTAGCGAGAGCAATTGATTGTTAAATTTGAGAGAAAGATCG 2670
QY 1027 TCTAGCTGATCAGGATCATATCTGTCGGGTCTTTTTCGGCTCAGTCAATCGCCCAAGCT 1086
Db |||||
2671 CGAGGAAGATCAATACATAAAGAGTTGAATCTTTTGTGTTCTTCGACATGGGTAACTCT 2730
QY 1087 GCGCTATCTGCGCATCGGGAGGAAGACCGCTGCTTTTCCCGCAGGTTGAAGCGG 1146
Db |||||
2731 CATGTTGNAATGCGCCTAGAGATCGGCGCAAGCTTGCAATGCTGCACTCTAGA 2790
QY 1147 CATGAAAGAGTTTGCCGAGGATGACTGCTGTCATTTGACGTTGAGCGTGAAGCAACGCACT 1206
Db |||||
2791 GGATCCCGACGCTCGACGCCATTAATAATGTTTTCGGTAAATTCAGCGCCTTCCATGAT 2850
QY 1207 TTACCATGATGATTCGGAGAGTGTTGGCCATCGACGCTTTTAAACGGTGAATCTGTTGTC 1266
Db |||||
2851 GAGACAGGCGGTTT--GAATGTTGAGCGGATGAACATATAAGCAATGACGGCAGCAATA 2908
QY 1267 AGGCCACCTGGGATACCACTGCTCGCGCTTTTCCGGACACAGTTCCGGATGCTCAGCG 1326
Db |||||
2909 AACTCAACAGG---AGCAGAAAGCGGGATATCCACAAAGTCCAGCGTACCAATAAGCG 2965
QY 1327 CGAAGCGCATCAGCAACCCGAAACATACCGGCGACAG-----CGGAACTGCCGTGCCG 1380
Db |||||
2966 CAAGCCTCAACGACGACGACGACGAGAGCGGTGAGTAGCAATCCAAATTTGTTACTC 3025
QY 1381 GTGTGAGATTAATGACACGCGGTGCGCGCTGGGATATTAATGCTCAGCGAGGACGGGTAT- 1439
Db |||||
3026 GTCAGAAAATCGAAATCATCTTTTCGTTAAATCCAAAACGGCAGAGCCCTGAAATGAGAA 3085
QY 1440 ---CTTGGCTGATGCGCGAGAAAATGGAATGATACCCCGTGAGTTACCGCGCGCGCG 1496
Db |||||
3086 CGACCTCGAGGGGGGCGCGTACCGAGCTTTTGTTCCTTTTAGTGAGGGTTAAATTCGA 3145
QY 1497 GCTTGGCGGTAATCATGCTCATAGCTGTTTCTGTTGAAATTTGTTATCCGCTCACAATTC 1556

3146 GCTTGGCGTAATCATGCTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATTC 3205
QY |||||
1557 CACACAAATACGACGCGGAGCATAAAGTGTAAAGCTTGGGTGCTTAATCAGTGAAGCT 1616
Db |||||
3206 CACACAAATAGGAGCCGGAAGCATAAAGTGTAAAGCTTGGGTGCTTAATCAGTGAAGT 3265
QY |||||
1617 AACTCACATTAATTCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGCTGCC 1676
Db |||||
3266 AACTCACATTAATTCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGCTGCC 3325
QY |||||
1677 AGCTGCAATTAATGAATTCGCGCAACCGCGGGGAGAGCGGTTTGGTATTTGGCGCTCTTT 1736
Db |||||
3326 AGCTGCAATTAATGAATTCGCGCAACCGCGGGGAGAGCGGTTTGGTATTTGGCGCTCTTT 3385
QY |||||
1737 CCGCTTCCCTCGCTCACTGACTCGCTCGCTCGCTGCTGCTGCTGCGGAGCGGATATCAG 1796
Db |||||
3386 CCGCTTCCCTCGCTCACTGACTCGCTCGCTCGCTCGCTGCTGCGGAGCGGATATCAG 3445
QY |||||
1797 CTCACCTCAAAGCGGTAATACGTTTATCCACAGAAATCAGGGGATAAACGAGGAAAGAAACA 1856
Db |||||
3446 CTCACCTCAAAGCGGTAATACGTTTATCCACAGAAATCAGGGGATAAACGAGGAAAGAAACA 3505
QY |||||
1857 TGTGACAAAAGGCGCAGCAAAAGCGCAGAAACCGTAAAGGCGCGGTTGCTGGCGCTTTT 1916
Db |||||
3506 TGTGACAAAAGGCGCAGCAAAAGGCGCAGAAACCGTAAAGGCGCGGTTGCTGGCGCTTTT 3565
QY |||||
1917 TCCATAGGCTCGCGCCCTGACGAGCATCAAAAATTCGACGCTCAAGTCAGAGGTGCG 1976
Db |||||
3566 TCCATAGGCTCGCGCCCTGACGAGCATCAAAAATTCGACGCTCAAGTCAGAGGTGCG 3625
QY |||||
1977 GAAACCCGACAGGACTATAAAGATACAGGCGCTTTTCCCGCTGGAAGCTCCCTCGCGCT 2036
Db |||||
3626 GAAACCCGACAGGACTATAAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGCGCT 3685
QY |||||
2037 CTCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTTCCCTTCGGGAAGCG 2096
Db |||||
3686 CTCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTTCCCTTCGGGAAGCG 3745
QY |||||
2097 TGGCGCTTTTTCATAGCTCAGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGTTCCGTCA 2156
Db |||||
3746 TGGCGCTTTTTCATAGCTCAGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGTTCCGTCA 3805
QY |||||
2157 AGCTGGGCTGTGTGACGAAACCCCGCTTCAGSCCGACGCGTTCGCGCTTTATCCGGTAACT 2216
Db |||||
3806 AGCTGGGCTGTGTGACGAAACCCCGCTTCAGSCCGACGCGTTCGCGCTTTATCCGGTAACT 3865
QY |||||
2217 ATCGCTTTGAGTCCAAACCGGTAAGACAGACTTATCGCCACTGCGCAGCAGCACTGGTA 2276
Db |||||
3866 ATCGCTTTGAGTCCAAACCGGTAAGACAGACTTATCGCCACTGCGCAGCAGCACTGGTA 3925
QY |||||
2277 ACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTTGAAGTGGTGGCCTA 2336
Db |||||
3926 ACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTTGAAGTGGTGGCCTA 3985
QY |||||
2337 ACTACGGCTTACACTAGAAGGACAGTATTTGGTATCTGGCGCTCTGCTGAAGCCAGTTACCT 2396
Db |||||
3986 ACTACGGCTTACACTAGAAGGACAGTATTTGGTATCTGGCGCTCTGCTGAAGCCAGTTACCT 4045
QY |||||
2397 TCGGAAAAAGAGTTGGTGAAGTTGATCCGGCAAAACAAACCCCGTGGTGGTGGTGGT 2456
Db |||||
4046 TCGGAAAAAGAGTTGGTGAAGTTGATCCGGCAAAACAAACCCCGTGGTGGTGGTGGT 4105
QY |||||
2457 TTTTGTGTTTGAAGCAGCAGATTTACCGCAGAAAAAAAGGATCTCAAGAAAGATCTTTTCA 2516
Db |||||
4106 TTTTGTGTTTGAAGCAGCAGATTTACCGCAGAAAAAAAGGATCTCAAGAAAGATCTTTTGA 4165
QY |||||
2517 TCTTTTCTACGGGCTCTGACGCTCAGTGAGAACGAAAACTCAGCTTAAAGGATTTTGGTCA 2576
Db |||||
4166 TCTTTTCTACGGGCTCTGACGCTCAGTGAGAACGAAAACTCAGCTTAAAGGATTTTGGTCA 4225
QY |||||
2577 TGAGATTAATCAAAAAGGATCTTCACTAGATCTTTTAAATTTAAAAATGAAGTTTAAAT 2636

Db 4226 TGAGATTATCAAAAGGATCTTCACCTAGATCCCTTTTAAATTAATAATGAAGTTTAAAT 4285

Qy 2637 CAATCTAAAGATATATATAGTAAACTTGTCTGACAGTTTACCAATGCTTAAATCAGTGAGG 2696

Db 4286 CAATCTAAAGTATATATAGTAAACTTGTCTGACAGTTTACCAATGCTTAAATCAGTGAGG 4345

Qy 2697 CACCTATCTCAGGATCTGTCTATTTTCTGCTATCCATAGTTCGCTGACTCCCGTCTGT 2756

Db 4346 CACCTATCTCAGGATCTGTCTATTTTCTGCTATCCATAGTTCGCTGACTCCCGTCTGT 4405

Qy 2757 AGATACTACGATACCGGAGGCTTACCATCTGGCCCGAGTCTGCAATGATACCCGAG 2816

Db 4406 AGATACTACGATACCGGAGGCTTACCATCTGGCCCGAGTCTGCAATGATACCCGAG 4465

Qy 2817 ACCACGCTCAGCGGCTCCAGATTTTATCAGCAATAAACACGAGCCGAGGCGGAGC 2876

Db 4466 ACCACGCTCAGCGGCTCCAGATTTTATCAGCAATAAACACGAGCCGAGGCGGAGC 4525

Qy 2877 GCAGAAGTGTCTGCAACTTTTATCCGCTCCATCCAGTCTATTATTTGTTGCCGGAG 2936

Db 4526 GCAGAAGTGTCTGCAACTTTTATCCGCTCCATCCAGTCTATTATTTGTTGCCGGAG 4585

Qy 2937 CTAGACTAGTGTCTGCGAGTTTAAAGTTTGGCAACGTTTGTGCAATGCTTACAGGCA 2996

Db 4586 CTAGACTAGTGTCTGCGAGTTTAAAGTTTGGCAACGTTTGTGCAATGCTTACAGGCA 4645

Qy 2997 TCGTGTGTACGCTCGTCTGTTGGTATGCTTTCATTCAGCTCCGTTCCCAACGATCAA 3056

Db 4646 TCGTGTGTACGCTCGTCTGTTGGTATGCTTTCATTCAGCTCCGTTCCCAACGATCAA 4705

Qy 3057 GGGAGTTACATGATCCCGATGTTGTGCAAAAAGCGTTAGTCTCTTCCGTTCTCCGA 3116

Db 4706 GGGAGTTACATGATCCCGATGTTGTGCAAAAAGCGTTAGTCTCTTCCGTTCTCCGA 4765

Qy 3117 TCGTGTCTCAGAGTAAAGTTGGCGAGTGTATCACTCATGTTTGGCAGCACTGCATA 3176

Db 4766 TCGTGTCTCAGAGTAAAGTTGGCGAGTGTATCACTCATGTTTGGCAGCACTGCATA 4825

Qy 3177 ATTCTCTTACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCA 3236

Db 4826 ATTCTCTTACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCA 4885

Qy 3237 AGTCACTCTGAGATAGTGTATGCGCGGACCGAGTTGCTCTTCCCGGCGTCAATACGGG 3296

Db 4886 AGTCACTCTGAGATAGTGTATGCGCGGACCGAGTTGCTCTTCCCGGCGTCAATACGGG 4945

Qy 3297 ATAATACCGCGCACATAGCAGAACTTTAAAGTGTCTCATTTGGAAACGTTCTTCGG 3356

Db 4946 ATAATACCGCGCACATAGCAGAACTTTAAAGTGTCTCATTTGGAAACGTTCTTCGG 5005

Qy 3357 GCGGAAACTCTCAAGGATCTTACCCTGTTCAGATCCAGTTCGATGTAAACCCACTCTGTG 3416

Db 5006 GCGGAAACTCTCAAGGATCTTACCCTGTTCAGATCCAGTTCGATGTAAACCCACTCTGTG 5065

Qy 3417 CACCCAACTGATCTTCAGATCTTTTATCTTTTACACAGGTTTCTGGGTGAGCAAAAACAG 3476

Db 5066 CACCCAACTGATCTTCAGATCTTTTATCTTTTACACAGGTTTCTGGGTGAGCAAAAACAG 5125

Qy 3477 GAAGCAAAATCCGCAAAAAGGGAATAAGGCGACAGGAAATGTTGAAATCTCATAC 3536

Db 5126 GAAGCAAAATCCGCAAAAAGGGAATAAGGCGACAGGAAATGTTGAAATCTCATAC 5185

Qy 3537 TCTTCTTTTCAATATTATGAAGCAATTTATCAGGTTTATGCTCTCATGAGCGGATACA 3596

Db 5186 TCTTCTTTTCAATATTATGAAGCAATTTATCAGGTTTATGCTCTCATGAGCGGATACA 5245

Qy 3597 TATTTGAATGTATTTAGAAAAATAAATAAGGGTTCCGCGCAATTTCCCGGAAAAAG 3656

Db 5246 TATTTGAATGTATTTAGAAAAATAAATAAGGGTTCCGCGCAATTTCCCGGAAAAAG 5305

Qy 3657 TGCCAC 3662

Db 5306 TGCCAC 5311

RESULT 8

I15649

LOCUS

DEFINITION Sequence 4 from patent US 5470573.

ACCESSION I15649

VERSION I15649.1

KEYWORDS GI:1250557

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 5314)

AUTHORS Lubitz,W. and Szostak,M.P.

TITLE Immunogens comprising the non-lytic membrane spanning domain of bacteriophages MS2 or PhiX174

JOURNAL Patent: US 5470573-A 4 28-NOV-1995;

FEATURES

source

1..5314

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match 71.6%; Score 2623.6; DB 6; Length 5314;

Best Local Similarity 83.4%; Pred.No. 0;

Mismatches 579; Indels 28; Gaps 6;

Matches 3059; Conservative 0;

QY 20 ATTTTGTAAAAATTCGCGTTAAATTTTGTAAATCAGCTCATTTTAAATCAATAGGCC 79

DB 1651 ATTTTGTAAAAATTCGCGTTAAATTTTGTAAATCAGCTCATTTTAAATCAATAGGCC 1710

QY 80 GAAATCGCAAAATCCCTTATAAATCAAAAGATAGACCGAGATAGGTTAGTGTGT 139

DB 1711 GAAATCGCAAAATCCCTTATAAATCAAAAGATAGACCGAGATAGGTTAGTGTGT 1770

QY 140 CCAGTTTGAACAAGAGTCCACTATTTAAAGACGTGACTCCAACTCAAGGGCGGAAA 199

DB 1771 CCAGTTTGAACAAGAGTCCACTATTTAAAGACGTGACTCCAACTCAAGGGCGGAAA 1830

QY 200 ACCGTCTATCAGGGCGATGGCCCACTATCGTGAACCATCACCTTAATCAAGTTTGGGG 259

DB 1831 ACCGTCTATCAGGGCGATGGCCCACTATCGTGAACCATCACCTTAATCAAGTTTGGGG 1890

QY 260 TCGAGTGTCCGTAAGACACTAAATCGAAACCTTAAAGGAGGCCCGATTTAGAGTTGA 319

DB 1891 TCGAGTGTCCGTAAGACACTAAATCGAAACCTTAAAGGAGGCCCGATTTAGAGTTGA 1950

QY 320 CCGGGAAGCCCGCGAACGTTGGCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 379

DB 1951 CCGGGAAGCCCGCGAACGTTGGCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2010

QY 380 AGGGCGCTGGCAAGTGTAGCGGTACGCTGCGGTAAACCAACACACCCCGCGCTTAAT 439

DB 2011 AGGGCGCTGGCAAGTGTAGCGGTACGCTGCGGTAAACCAACACACCCCGCGCTTAAT 2070

QY 440 GCGCCGCTACAGGGCGGCTCCCATTCGCCATTGAGCTCGCAACTGTGTGGGAAGGCGGA 499

DB 2071 GCGCCGCTACAGGGCGGCTCCCATTCGCCATTGAGCTCGCAACTGTGTGGGAAGGCGGA 2130

QY 500 TCGGTGCGGGCTCTTCGCTATTTACCCAGCTGGCGAAGGAGGAGGAGGAGGAGGAGGAG 559

DB 2131 TCGGTGCGGGCTCTTCGCTATTTACCCAGCTGGCGAAGGAGGAGGAGGAGGAGGAGGAG 2190

QY 560 TTAAGTTGGTTAAACCGCAGGGTTTCCAGTCAAGCTTGTAAAAACGACGCGCAGTGAG 619

DB 2191 TTAAGTTGGTTAAACCGCAGGGTTTCCAGTCAAGCTTGTAAAAACGACGCGCAGTGAA 2250

QY 620 CCGCGCTCGTTCAATTCAGCTTTTGAACCCGCTGGAGGACGCGCAGACTCGCGGTGCAAT 679

DB 2251 TTGTAATACGACTCACTATAGGCGGAATTTGGAGCTCCACCGCGGTGGCGCGCTCTAGT 2310

QY 680 GTGTTTACGCGTGTAGTGAGCAGATGAAGATGCTCGACACCGTGCAGAACACGCGACTA 739

Dd 2311 ATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGATATTTAAGCCAGTATATACACTC 2370
Qy 740 GATTAAACCTCAGAAAGATAATCATATTGTGAGCTAGCTTTAAAGATAATCATCGTAAAT 799
Dd 2371 CGCTATCGTACGTGGTCATGCGCTGCGCCCGCACCCGCGCAACCCCGTGACG 2430
Qy 800 TGACGCATGGGATCTGTAAATAGCACTCACTATAGGGCGAATTTGGGTACCGGGCCCCCCT 859
Dd 2431 CGCCCTGACGGCTGTCTGCTCCCGCATCGCTTACAGACAAGCTGTGACGCTCTCG 2490
Qy 860 CGAGGTGAGCGTATCGA-----TAAGCTTGATATCGAATTCCTGACGCCCGGGGATC 913
Dd 2491 GGAGTGCATGTGTGTCAGAGGTTTTCACCGTCAATCACCGAAACCGCGAGCGAGTAAAGTTC 2550
Qy 914 CACTAGTTCTAGAGGGCGCCACCGCGGTGAGCTCCAGCTTTTGTCCCTTTAGTGAG 973
Dd 2551 GGATGCTTTGTGAGCAATTCGTCCTTAAAGTAAGCAATTCGTGTAAGTGTCACTGTGC 2610
Qy 974 GGTAAATTAGATCCCAT-----GCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAA 1026
Dd 2611 GGATCACCGCTCCAGTAGCGAGAGAAATTTGATTGGTAAATTTGAGAGAAAGATCG 2670
Qy 1027 TCTAGTGCATCAGGATCATATCTGTCGGTCTTTTTCGGCTCAGTCAATCCCCAAAGCT 1086
Dd 2671 CGAGGAAGATCAATACATAAAGAGTTGAATCTTTTGTGTCTTCGACATGGGTAAATCCT 2730
Qy 1087 GCGCTATCTGGGCATCGGGAGGAAGAGCCGTGCTTTTCCGCGAGGTTGAAGCG 1146
Dd 2731 CATGTTTGAATGCGCCTAGAGATCGGCGCAAGCTTGATGCTGCGAGTCACTCTAGA 2790
Qy 1147 CATGGAAGAGTTTGC CGAGGATGACTGCTGCTGCAATTGACGTTGAGCGAAGCAACGAT 1206
Dd 2791 GGATCCCGACGCTCGACGCCAATAATATGTTTTCGGTAAATTCAGCGCCTTCATGAT 2850
Qy 1207 TTACCATGATGATTCGGGAAGTGTGGCCATCGACGCCCTTTAAACGGTGAATGTTGCTTC 1266
Dd 2851 GAGACAGGCGGTTT--GAATGTTGAGGGATGAACAATAAAGCAATGACGCGACGAATA 2908
Qy 1267 AGGCCACCTGGGATACCACTGCTGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCC 1326
Dd 2909 AACTCAACAGG---AGCAGGAAGCGAGGATATCCACAAGTCCAGCGTACCATAAAGC 2965
Qy 1327 CGAAGCGCATCAGCAACCCGGAACAATAACCGGACAG-----CCGGAACCTCCGTCGCG 1380
Dd 2966 CAAGCTCAACGACGACGACGACGAGAGCGGTGAGTACGAATCCAAACTTTGTACTC 3025
Qy 1381 GTGTGAGATTAATGACAGCGGTGGCGCTGGGATATTAAGTCAAGCGAGGACGGGTAT- 1439
Dd 3026 GTCAGAAAATCGAAAATCATCTTCGTTAAATCCAAAACGGCAGAGACCTGAATGAGAAAT 3085
Qy 1440 ---CCTGGCTGATGCCGAGAAATGGACATGGATACCCCGTGGATTACC CGCGCGCGC 1496
Dd 3086 CGACCTCGAGGGGGGCGCGTACCAGCTTTTGTCCCTTTAGTGAGGTTAAATCCGA 3145
Qy 1497 GCTTGGCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTC 1556
Dd 3146 GCTTGGCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTC 3205
Qy 1557 GACACAACTACGACCGCGAAGCATAAAGTGAAGCCTGGGGTGGCTTAATCAGTGAGCT 1616
Dd 3206 CACACAACATAGGACCGGAAGCATAAAGTGAAGCCTGGGGTGGCTTAATCAGTGAGCT 3265
Qy 1617 AACTCACATTAATTTGGTTGGCTCACTGCCGCTTTCCAGTCCGGAACCTGTGCTGCC 1676
Dd 3266 AACTCACATTAATTTGGTTGGCTCACTGCCGCTTTCCAGTCCGGAACCTGTGCTGCC 3325
Qy 1677 AGTCGATTAATGAATTCGGCCAAACCGCGGGGAGAGCGGTTTGTGATATTGGCGCTCTT 1736
Dd 3326 AGTCGATTAATGAATTCGGCCAAACCGCGGGGAGAGCGGTTTGTGATATTGGCGCTCTT 3385
Qy 1737 CCGCTTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 1796
Dd 3386 CCGCTTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 3445

Qy 1797 CTCACTCAAAGCGGTAATAACGGTTTATCCACAGAAATCAGGGGATAAACGAGGAAGAAACA 1856
Dd 3446 CTCACTCAAAGCGGTAATAACGGTTTATCCACAGAAATCAGGGGATAAACGAGGAAGAAACA 3505
Qy 1857 TGTGAGCAAAAAGCGCAGCAAAAAGCGCAGAAACCGTAAAAAAGGCCGCTGCTGGCGTTTT 1916
Dd 3506 TGTGAGCAAAAAGCGCAGCAAAAAGCGCAGAAACCGTAAAAAAGGCCGCTGCTGGCGTTTT 3565
Qy 1917 TCCATAGGCTCGGCCCGCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGCG 1976
Dd 3566 TCCATAGGCTCGGCCCGCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGCG 3625
Qy 1977 GAAAACCCGACAGCACTATAAAGATACAGAGCGCTTTCCGCCCTGGAAGCTCCCTCGTCGCT 2036
Dd 3626 GAAAACCCGACAGCACTATAAAGATACAGAGCGCTTTCCGCCCTGGAAGCTCCCTCGTCGCT 3685
Qy 2037 CTCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCG 2096
Dd 3686 CTCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCG 3745
Qy 2097 TGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGCTGAGGTGCTTCGCTCCA 2156
Dd 3746 TGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGCTGAGGTGCTTCGCTCCA 3805
Qy 2157 AGCTGGGCTGTGTGACGAAACCCCGCTTTCAGCCCGACCGCTGCGCTTATCCGTTAACT 2216
Dd 3806 AGCTGGGCTGTGTGACGAAACCCCGCTTTCAGCCCGACCGCTGCGCTTATCCGTTAACT 3865
Qy 2217 ATCGCTTTGAGTCCAAACCGGTAAAGACAGCTTATCGCACCTGCGCAGCAGCACTGGTA 2276
Dd 3866 ATCGCTTTGAGTCCAAACCGGTAAAGACAGCTTATCGCACCTGCGCAGCAGCACTGGTA 3925
Qy 2277 ACAGATTTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTA 2336
Dd 3926 ACAGATTTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTA 3985
Qy 2337 ACTACGGCTTACTAGAAAGGACAGTATTTGGTATCTGCGCTCTGCTGAGCAGCTTAACT 2396
Dd 3986 ACTACGGCTTACTAGAAAGGACAGTATTTGGTATCTGCGCTCTGCTGAGCAGCTTAACT 4045
Qy 2397 TCGGAAAAAGAGTGGTGTAGCTTGTATCCGCAAAACAAACCCGCTGCTGAGCGGTGTT 2456
Dd 4046 TCGGAAAAAGAGTGGTGTAGCTTGTATCCGCAAAACAAACCCGCTGCTGAGCGGTGTT 4105
Qy 2457 TTTTGTGTTGCAAGCAGCAGATTACGCGAGAAAAAAGAGATCTCAAGAGATCTCTTGA 2516
Dd 4106 TTTTGTGTTGCAAGCAGCAGATTACGCGAGAAAAAAGAGATCTCAAGAGATCTCTTGA 4165
Qy 2517 TCTTTTCTACGGGCTGACGCTCAGTGGAAACGAAAACTCAGCTTAAAGGATTTTGGTCA 2576
Dd 4166 TCTTTTCTACGGGCTGACGCTCAGTGGAAACGAAAACTCAGCTTAAAGGATTTTGGTCA 4225
Qy 2577 TGAGATTAATAAAAAGGATCTTCACTAGATCTTTTAAATTTAAAAATGAAGTTTAAAT 2636
Dd 4226 TGAGATTAATAAAAAGGATCTTCACTAGATCTTTTAAATTTAAAAATGAAGTTTAAAT 4285
Qy 2637 CAATCTAAAGTATATATAGTAAATCTGCTGACAGTTTACCAATGCTTAAATCAGTGAG 2696
Dd 4286 CAATCTAAAGTATATATAGTAAATCTGCTGACAGTTTACCAATGCTTAAATCAGTGAG 4345
Qy 2697 CACCTATCTCAGCGATCTGTCTATTTTGTTCATTCATAGTTGCTGACTCCCGCTGCT 2756
Dd 4346 CACCTATCTCAGCGATCTGTCTATTTTGTTCATTCATAGTTGCTGACTCCCGCTGCT 4405
Qy 2757 AGATAACTACGATACGGGAGGCTTACCATCTGGGCCCGAGTCTGCAATGATACCGGAG 2816
Dd 4406 AGATAACTACGATACGGGAGGCTTACCATCTGGGCCCGAGTCTGCAATGATACCGGAG 4465
Qy 2817 ACCCAGCTCACCGGCTCCAGATTTTATCAGCAATAAACCAGCAGCCGGAAGGCGGAGC 2876
Dd 4466 ACCCAGCTCACCGGCTCCAGATTTTATCAGCAATAAACCAGCAGCCGGAAGGCGGAGC 4525

QY 2877 GCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTATTTGTTGCGGGAAG 2936
Db 4526 GCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTATTTGTTGCGGGAAG 4585
QY 2937 CTAGAGTAAAGTAGTTCGCGCAAGTTAAAGTTTGGCAACGTTGTTGCACTTGTCTACAGGCA 2996
Db 4586 CTAGAGTAAAGTAGTTCGCGCAAGTTAAAGTTTGGCAACGTTGTTGCACTTGTCTACAGGCA 4645
QY 2997 TGTGTGTACAGCTCGTCTGTTGGTATAGCTTTTATTCAGTCTCGGTTCCCAACGATCAA 3056
Db 4646 TGTGTGTACAGCTCGTCTGTTGGTATAGCTTTTATTCAGTCTCGGTTCCCAACGATCAA 4705
QY 3057 GCGAGTTACATCATCCCCCATGTTGTCGCAAAAAGGTTAGTCTTGGTCTCCGCA 3116
Db 4706 GCGAGTTACATCATCCCCCATGTTGTCGCAAAAAGGTTAGTCTTGGTCTCCGCA 4765
QY 3117 TGTGTCTCAGAAAGTAAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTCATTA 3176
Db 4766 TGTGTCTCAGAAAGTAAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTCATTA 4825
QY 3177 ATTCTCTTACTGTATGCCATCGTAAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCA 3236
Db 4826 ATTCTCTTACTGTATGCCATCGTAAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCA 4885
QY 3237 AGTCATTCTGAGAAATAGTATGCGCGCAGGTTGCTCTTCCCGCGGTCAATACGGG 3296
Db 4886 AGTCATTCTGAGAAATAGTATGCGCGCAGGTTGCTCTTCCCGCGGTCAATACGGG 4945
QY 3297 ATAAATACCGCGCACATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAACGTTCTTCGG 3356
Db 4946 ATAAATACCGCGCACATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAACGTTCTTCGG 5005
QY 3357 GCGCAAACTCTCAAGGATCTTACCGCTGTGTAGATCCAGTGTGATGTAACCCACTCGTG 3416
Db 5006 GCGCAAACTCTCAAGGATCTTACCGCTGTGTAGATCCAGTGTGATGTAACCCACTCGTG 5065
QY 3417 CACCCAACTGTCTCAGGATCTTTTACCTTTCACAGCGTTTCTGGTGAGCAAAAACAG 3476
Db 5066 CACCCAACTGTCTCAGGATCTTTTACCTTTCACAGCGTTTCTGGTGAGCAAAAACAG 5125
QY 3477 GAAGGCAAAATGCGCAAAAAGGGAATAAGGCGCACAGGAAATGTTCAATCTCATAC 3536
Db 5126 GAAGGCAAAATGCGCAAAAAGGGAATAAGGCGCACAGGAAATGTTCAATCTCATAC 5185
QY 3537 TCTTCCTTTTCAATATTTAAGCAATTTATCAGGGTTTATGTCATCAGCGGATACA 3596
Db 5186 TCTTCCTTTTCAATATTTAAGCAATTTATCAGGGTTTATGTCATCAGCGGATACA 5245
QY 3597 TATTTGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGGCGACATTTCCCGAAAAAG 3656
Db 5246 TATTTGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGGCGACATTTCCCGAAAAAG 5305
QY 3657 TGCCAC 3662
Db 5306 TGCCAC 5311

RESULT 9
AF171227/c
LOCUS AF171227 3552 bp DNA circular SYN 04-FEB-2000
DEFINITION Shuttle vector pSLIRES11, complete sequence.
ACCESSION AF171227
VERSION AF171227.1 GI:5764544
KEYWORDS
SOURCE Shuttle vector pSLIRES11
ORGANISM Shuttle vector pSLIRES11
other sequences; artificial sequences; vectors.
REFERENCE 1 (bases 1 to 3552)
AUTHORS Chen, C.M., Smith, D.M., Peters, M.A., Samson, M.E., Zitz, J.,
Tabin, C.J., and Cepko, C.L.
TITLE Production and design of more effective avian
replication-incompetent retroviral vectors
JOURNAL Dev. Biol. 214 (2), 370-384 (1999)

MEDLINE 99456779
PUBMED 10525341
REFERENCE 2 (bases 1 to 3552)
AUTHORS Chen, C.-M.A., Samson, M.E.S. and Cepko, C.L.
TITLE Direct Submision
JOURNAL Submitted (23-JUL-1999) Genetics, Harvard Medical School, 200
Longwood Avenue, Boston, MA 02115, USA
FEATURES
source
1..3552
/organism="Shuttle vector pSLIRES11"
/mol_type="genomic DNA"
/db_xref="taxon:101936"
/lab_host="Escherichia coli"
join(2864..3552,1..2208)
/note="similar to pBluescript"
2236..2789
/note="IRES of EMCV"
2790..2849
/note="multiple cloning site"
ORIGIN
Query Match 71.3%; Score 2610.8; DB 12; Length 3552;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 482; Indels 110; Gaps 6;
QY 1 CTAAATGTTAAGCGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTC 60
Db 3552 CTAAATGTTAAGCGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTC 3493
QY 61 ATTTTAAACCAATAGCGCAAAATCGGCAAAATCCCTTAAATCAAAAGAAATAGACCGA 120
Db 3492 ATTTTAAACCAATAGCGCAAAATCGGCAAAATCCCTTAAATCAAAAGAAATAGACCGA 3433
QY 121 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGGACTC 180
Db 3432 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGGACTC 3373
QY 181 CAACGTCAAAGGCGCAAAACCGTCTATCAGGCGATGCGCCACTACGTGAACCATCACC 240
Db 3372 CAACGTCAAAGGCGCAAAACCGTCTATCAGGCGATGCGCCACTACGTGAACCATCACC 3313
QY 241 CTAAATCAAGTTTGGGTCGAGTCCGTAAGCACTTAAATCGGAACCTTAAAGGGAG 300
Db 3312 CTAAATCAAGTTTGGGTCGAGTCCGTAAGCACTTAAATCGGAACCTTAAAGGGAG 3253
QY 301 CCCCCGATTTAGAGCTTTGACGCGGAAAGCCGCGAAGCTGGCGAGAAAGGAAGGAAGAA 360
Db 3252 CCCCCGATTTAGAGCTTTGACGCGGAAAGCCGCGAAGCTGGCGAGAAAGGAAGGAAGAA 3193
QY 361 AGCGAAAGAGCGGCGCTTAGGCGCTGGCAAGTGTAGCGGTACCGCTGCGGTAAACCAC 420
Db 3192 AGCGAAAGAGCGGCGCTTAGGCGCTGGCAAGTGTAGCGGTACCGCTGCGGTAAACCAC 3133
QY 421 CACACCCGCGCTTAAATGCGCGCTACAGGCGGCTCCCATTCGCCATTTCAGGCTGGC 480
Db 3132 CACACCCGCGCTTAAATGCGCGCTACAGGCGGCTCCCATTCGCCATTTCAGGCTGGC 3073
QY 481 CAACTGTTGGGAAGGCGGATCGGTGCGGCTCTTCGCTATTACGCGAGCTGGCGAAAGG 540
Db 3072 CAACTGTTGGGAAGGCGGATCGGTGCGGCTCTTCGCTATTACGCGAGCTGGCGAAAGG 3013
QY 541 GGGATGTGCTCAAGCGGATTAAGTTGGGTAAAGCGCAGGGTTTCCAGTACGAGCTTG 600
Db 3012 GGGATGTGCTCAAGCGGATTAAGTTGGGTAAAGCGCAGGGTTTCCAGTACGAGCTTG 2953
QY 601 TAAACGAGCGGCGAGTACGCGGCTCTGCTTTCATTCAGTTTGTGAACCCGCTGAGACGG 660
Db 2952 TAAACGAGCGGCGAGTACGCGGCTCTGCTTTCATTCAGTTTGTGAACCCGCTGAGACGG 2913
QY 661 GCAGACTCGCGTGCATAATGTGTTTACAGCGTGTAGGAGCAGATGAAGATGCTCGACAC 720
Db 2912 GCGGAATGGGTACCGGGCCCCCTCGATCGAGGTGCGAGGTATCGATAGCTTGGGCT 2853

QY 721 GCTGCAGAACCGCAGCTAGATTAAACCCCTAGAAAAGATAATCATATTGTGACGTACGTTAA 780
Db 2852 GCAGCG-----CCGCACTAGTGATCGGATCCCGGCGAGCTCGAAT 2808
QY 781 AGATAATCATGCGTAAATTTGACGATGGGATCTGTATAACGACTCATATAGGCGCGAAT 840
Db 2807 TCGTAATCATGCGCATGCGTATTATCATCGTGTGTTTTTCAAAGGAAAACCAACGTCCTCCCGTGG 2748
QY 841 TGGGTACCGGGCCCCCTCGAGGTCGACGGTATCGATAAGCTTCGATATCGAATTCCTGC 900
Db 2747 TTCGGGGGCCCTAGACGTTTTTAACTCGACTAAACACACGTAAGCATGTGACCGAG 2688
QY 901 AGCCCGGGGGATCCACTAGTTCTTAGAGCGCGCCACCGCGGTGGAGCTCCAGCTTTTTGT 960
Db 2687 GCCCCAGATCAGATCCCATACAATGGGTACCTTCGGGCATCCTTCAGCCCTTTGTTGA 2628
QY 961 TCCCTTTAGTAGGGTTAAATAGATCCCATGGTCAATTTTACGGCAGACTATCTTTCTAG 1020
Db 2627 ATACGCTTTGAGGAGAGCCATTTGACTCTTTCCACAACATATCCAACCTCAACAGTGGCACT 2568
QY 1021 GGTTAATCTAGCTGATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCC 1080
Db 2567 GG-----GGTTGCGCCCTTTGCGAGTGTATCTTATAACGTGGCTTTTGCC 2519
QY 1081 CAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAACCCGCTGCTTTTTCCCGCGAGTTG 1140
Db 2518 GCAGAGCACCTGTCCGAGGTGGGGTTCCGCTGCCCTGCAAGGGTCCGTACAGACGT 2459
QY 1141 AAGCGCATGGAAGAGTTTGCAGAGGATGACTGCTGCTGCATTTGACGTTGAGCGAAAC 1200
Db 2458 TGTTTTGTCTTCAAGAAGCTTCCAGAGGA-----ACTGCTCTCCTCAGCACATTCACAGA 2404
QY 1201 GCACGTTTACCATGATGATTCGGGAAGTGTGGCCATGCAGCCCTTTAAGGTTGAAGTGT 1260
Db 2403 CTTGCATTTCTTTGCGAGAGGGAAGAACCCCTAGGAATGCTCGTCAAGAGACAGG 2344
QY 1261 TCGTTTCAGGCCACCTGGGATACAGTTTCTGTCGGCTTTTCCGGACACAGTTTCCGGATGG 1320
Db 2343 CCAGGTTTCCGGGCCCTCACATTTGCAAAAGACGGCAATATGGTGGAAAATAACATATAG 2284
QY 1321 TCAGCCGAGAGCGCATCAGCAACCCGAAACATACCGGCGACAGCCGGAACTGCCGTGCG 1380
Db 2283 ACAACGCA-----CACCGCCCTTATTCCAAGCGGCTTCGGCCAGTAAGCTTAGGGGG 2231
QY 1381 GTGTGCAGATTAAACACAGCGGTGCGGCTGGGATATTACGTACGCGAGGACGGGTATC 1440
Db 2230 GGGGAGGAGTCTAGATCATGATAGCTCCAGCTTTTGTTCCTTTAGTGGGTTAAT- 2172
QY 1441 CTGGCTGGATGCCGAGAAATGGACATGGATACCCCGTGAGTTACCCGCGCGCGCTT 1500
Db 2171 -----TTCGAGCTT 2163
QY 1501 GCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATTCACA 1560
Db 2162 GCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATTCACA 2103
QY 1561 CAACATACAGCCGGAAGCATAAAGTGTAAAGCTTGGGTGCTTAATGAGTGAGCTAACT 1620
Db 2102 CAACATACAGCCGGAAGCATAAAGTGTAAAGCTTGGGTGCTTAATGAGTGAGCTAACT 2043
QY 1621 CACATTAAATGGTTGCGTCTCATCTGCCGCTTTCCAGTCGGGAAACCTGTGCGCCAGCT 1680
Db 2042 CACATTAAATGGTTGCGTCTCATCTGCCGCTTTCCAGTCGGGAAACCTGTGCGCCAGCT 1983
QY 1681 GCATTAAATGAATCGGCCAAACGCGCGGGGAGAGGCGGTTTGGTATTTGGCGCTCTTCGC 1740
Db 1982 GCATTAAATGAATCGGCCAAACGCGCGGGGAGAGGCGGTTTGGTATTTGGCGCTCTTCGC 1923
QY 1741 TTCCTCGCTCACTGACTCGCTGCGCTCGGTGCTGCGGCGAGCGGTATCAGCTCA 1800
Db 1922 TTCCTCGCTCACTGACTCGCTGCGCTCGGTGCTGCGGCGAGCGGTATCAGCTCA 1863
QY 1801 CTCAAAGCGGTAATACGGTTTATCCAGAAATCAGGGGATTAACGCGAGGAAACATGTG 1860

Db 1862 CTCAAAGCGGTAATACGGTTTATCCAGAAATCAGGGGATTAACGCGAGGAAACATGTG 1803
QY 1861 AGCAAAAGCGCAGCAAAAGGCGCAGGAACCGTAAAGGCGCGGTTGCTGGCGTTTTTCCA 1920
Db 1802 AGCAAAAGCGCAGCAAAAGGCGCAGGAACCGTAAAGGCGCGGTTGCTGGCGTTTTTCCA 1743
QY 1921 TAGGCTCCGCCCTGACGAGCATCAGAAAATCGAGCTCAAGTCAGAGGTGCGGAAA 1980
Db 1742 TAGGCTCCGCCCTGACGAGCATCAGAAAATCGAGCTCAAGTCAGAGGTGCGGAAA 1683
QY 1981 CCCGACAGGATATATAAGATACACAGCGGTTTTCCCTCTGGAAGCTCCCTCGTGGCTCTCC 2040
Db 1682 CCCGACAGGATATATAAGATACACAGCGGTTTTCCCTCTGGAAGCTCCCTCGTGGCTCTCC 1623
QY 2041 TGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGC 2100
Db 1622 TGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGC 1563
QY 2101 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTCTGCTCCAAGCT 2160
Db 1562 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTCTGCTCCAAGCT 1503
QY 2161 GGGCTGTGTGACGAAACCCCGCTTTCAGCCGACCGCTGCGCTTATCCGGTAACTATCG 2220
Db 1502 GGGCTGTGTGACGAAACCCCGCTTTCAGCCGACCGCTGCGCTTATCCGGTAACTATCG 1443
QY 2221 TCTTAGGTCCAACCCGGTAAGACAGACTTATCGCCACTTGGCAGCAGCCACTGGTAAACAG 2280
Db 1442 TCTTAGGTCCAACCCGGTAAGACAGACTTATCGCCACTTGGCAGCAGCCACTGGTAAACAG 1383
QY 2281 GATTAGCAGAGCGAGGTATGTAGCGGTGTACAGAGTTCTTGAAGTGGTGGCTTAACTA 2340
Db 1382 GATTAGCAGAGCGAGGTATGTAGCGGTGTACAGAGTTCTTGAAGTGGTGGCTTAACTA 1323
QY 2341 CGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGGTTTACCTTGG 2400
Db 1322 CGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGGTTTACCTTGG 1263
QY 2401 AAAAAGAGTTGTTAGTCTTGTATCCGCCAAACAAACACCGCTGTAGCGGTGGTTTTTT 2460
Db 1262 AAAAAGAGTTGTTAGTCTTGTATCCGCCAAACAAACACCGCTGTAGCGGTGGTTTTTT 1203
QY 2461 TGTTTGAAGCAGCAGATTCAGCGCAGAAAAAAGGATCTCAAGAAAGTCTTTCATCTT 2520
Db 1202 TGTTTGAAGCAGCAGATTCAGCGCAGAAAAAAGGATCTCAAGAAAGTCTTTCATCTT 1143
QY 2521 TTCTACGGGCTGTAGCGCTCAGTGGAAAGAAAACTCAGTTAAGGGATTTTGGTCATGAG 2580
Db 1142 TTCTACGGGCTGTAGCGCTCAGTGGAAAGAAAACTCAGTTAAGGGATTTTGGTCATGAG 1083
QY 2581 ATTATCAAAAAGGATCTTTCACCTAGATCTTTAAATTTAAAAATGAAGTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGGATCTTTCACCTAGATCTTTAAATTTAAAAATGAAGTTTAAATCAAT 1023
QY 2641 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTTACCAATGCTTAACTCAGTGAGGCA 2700
Db 1022 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTTACCAATGCTTAACTCAGTGAGGCA 963
QY 2701 TATCTCAGCGATCTGTCTATTCTTGGTTTCATAGTTCGCTGACTCCCGCTGCTGTAGAT 2760
Db 962 TATCTCAGCGATCTGTCTATTCTTGGTTTCATAGTTCGCTGACTCCCGCTGCTGTAGAT 903
QY 2761 AACTACGATACGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCCGAGAGACC 2820
Db 902 AACTACGATACGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCCGAGAGACC 843
QY 2821 ACCTCACCGCTCCAGATTTTATCAGCAATAAACACAGCAGCCGGAAGCGCGAGCGCAG 2880
Db 842 ACCTCACCGCTCCAGATTTTATCAGCAATAAACACAGCAGCCGGAAGCGCGAGCGCAG 783
QY 2881 AAGTGTCTCTGCACTTTATCGGCTCCATTCAGTCTATTATTTGTTCCGCGGAAGCTT 2940

Db	782	AGTGGTCTCGCAACTTTATCCGGCTCCATCCAGCTCTATTAAATTTGTTGCGGGAAGCTAG	723
Qy	2941	AGTAAGTAGTTTCGGCAGTTAATAGTTTGGCGAAAGTTTGGCCATTCGTACAGGCATCGT	3060
Db	722	AGTAAGTAGTTTCGGCAGTTAATAGTTTGGCGAAAGTTTGGCCATTCGTACAGGCATCGT	663
Qy	3001	GGTGTACAGCTTCGTTCGTTTGGTATGCGTTCATTCAGCTTCGGTTCCTCAACGATCAAGGCG	3060
Db	662	GGTGTACAGCTTCGTTCGTTTGGTATGCGTTCATTCAGCTTCGGTTCCTCAACGATCAAGGCG	603
Qy	3061	AGTTACATGATCCCCCATGTTTGCAGAAAGCGGTTAGCTTCCTTCGGTCTCTCCGATCGT	3120
Db	602	AGTTACATGATCCCCCATGTTTGCAGAAAGCGGTTAGCTTCCTTCGGTCTCTCCGATCGT	543
Qy	3121	TGTGAGAAGTAAGTTTGGCCGAGCTGTTTATCACTCATGCTTATGCGCAGCACTGCATAATTC	3180
Db	542	TGTGAGAAGTAAGTTTGGCCGAGCTGTTTATCACTCATGCTTATGCGCAGCACTGCATAATTC	483
Qy	3181	TCTTACTGTCTATGTCATCCGTAAGATGCTTTTCTGTGACTGTGTGAGTACTCAACAAAGTC	3240
Db	482	TCTTACTGTCTATGTCATCCGTAAGATGCTTTTCTGTGACTGTGTGAGTACTCAACAAAGTC	423
Qy	3241	ATTTCTGAGATAGTATGCGCGACCGAGTTGCTCTTTCGCCGCGCTCAATACGGGATAA	3300
Db	422	ATTTCTGAGATAGTATGCGCGACCGAGTTGCTCTTTCGCCGCGCTCAATACGGGATAA	363
Qy	3301	TACCGCGCCACATAGCAGAACTTTTAAAGTCTCATCATTTGGAAGAGTTTCTTCGGGGCG	3360
Db	362	TACCGCGCCACATAGCAGAACTTTTAAAGTCTCATCATTTGGAAGAGTTTCTTCGGGGCG	303
Qy	3361	AAAACTCTCAAGGATCTTACCGCTGTTTGAGATCCAGTTTCGANTGAAACCCACTCGTGCACC	3420
Db	302	AAAACTCTCAAGGATCTTACCGCTGTTTGAGATCCAGTTTCGANTGAAACCCACTCGTGCACC	243
Qy	3421	CAACTGATCTTCAGCATCTTTTACTTTTCCACAGCGTTTCTGGGTGAGCAAAAACAGGAAG	3480
Db	242	CAACTGATCTTCAGCATCTTTTACTTTTCCACAGCGTTTCTGGGTGAGCAAAAACAGGAAG	183
Qy	3481	GCAAAATGCCGCAAAAAGGAAATAAGGGCGACACGGAATGTTTGAATCTCATACTCTTT	3540
Db	182	GCAAAATGCCGCAAAAAGGAAATAAGGGCGACACGGAATGTTTGAATCTCATACTCTTT	123
Qy	3541	CTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTTCTCATGAGCGGATACATATT	3600
Db	122	CTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTTCTCATGAGCGGATACATATT	63
Qy	3601	TGAATGATTTTGAAGAAATAAACAATAGGGTTTCGGCGACATTTCCCGGAAAAGTGC	3660
Db	62	TGAATGATTTTGAAGAAATAAACAATAGGGTTTCGGCGACATTTCCCGGAAAAGTGC	3
Qy	3661	AC 3662	
Db	2	AC 1	
RESULT 10			
LOCUS	AX306330	4001 bp	DNA linear PAT 11-DEC-2001
DEFINITION	Sequence 13 from Patent WO0188121.		
ACCESSION	AX306330		
VERSION	AX306330.1	GI:17645569	
KEYWORDS	synthetic construct		
SOURCE	synthetic construct		
ORGANISM	other sequences; artificial sequences.		
REFERENCE	1		
AUTHORS	Plaetinck, G., Renard, J. P. and Bogaert, T.		
TITLE	Vector constructs		
JOURNAL	Patent: WO 0188121-A 13 22-NOV-2001;		
DEVEN	Devgen NV (BE)		
FEATURES	Location/Qualifiers		
source	1..4001		
	/organism="synthetic construct"		

Qy	933	GCACCGCGGTGGAGCTCCAGCTCTTTTGTTCCTCTTTAGTGAGGGTTAATTAGATCCCATGC	992
Db	1022	TGGAAGCGGAAAATCAGGAAGGATGGCTAGGTCGCCCGGTTTATTGGAATGAACGGC	1081
Qy	993	GTCAATTTTACGAGACTA--TCTTCTAGGGTTAATCTAGCTGCATCAGGATCATATC	1049
Db	1082	TCTTTTCTGACGAGAAACAGGACTGGTGAATATGCAAGTTTAAAGTTTACACCTATAAAAG	1141
Qy	1050	GTCCGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGGAG	1109
Db	1142	AGAGAGCGGTATCGTCTGTTTGTGGATGTACAGAGTGATATTATTTCACACGCCCGGGCG	1201
Qy	1110	GAAGAACCCGTGCTTTTCCCGGAGGTTGAAGCGGCATGGAAGAGTTTGGCCGAGGAT	1169
Db	1202	ACGGATGGTGATCCCCCTGGCC-AGTGCAAGTCTCTTAAAGCGATAAAGTCTCCCGTGAAC	1260
Qy	1170	GACTGCTGCTG--CATTTGACGTTTGAGCGAANAACGACGTTTACCATGATGATTTCCGGGAAG	1227
Db	1261	TTTACC CGGTGGTGATATCGGGGATGAAGAGCTGGCGCATGTAGCACCGATATGGCCA	1320
Qy	1228	GTGTGGCCATGACAGCCCTTTAAACGTTGAACTGTTTTCAGGCCACCTGGGATACAGATT	1287
Db	1321	GTGTGCCGCTCTCGTTTATCGGGGAGAGTGGCTGATCTCAGGCCACCGGAAAATGACA	1380
Qy	1288	CGTCCGGCTTTTCCGACACAGTTCCGGATGTCAGCCCGCAAGCGCATCAGCAACCCGA	1347
Db	1381	TCAAAACCGCATTTAACCTGATGTTCTTGGGGAATATAAATGTCAAGGCTCCCTTTATACA	1440
Qy	1348	ACAATACCGGGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACACGGTGGCG	1407
Db	1441	GCCTTTCCAGCACAAATGGATCTCGAGGGATCTTTCCATACCTACCGATTCTGCGCCTGCAG	1500
Qy	1408	CGCTGGGATATTACGTACAGCGAGACGGGTATCTCTGCTGGATGCCGCAAAATGACAT	1467
Db	1501	GTCCGCCCGGACTCTCTAGATGCGAAAGCTTCTCGCCCTATAGTGAGTCGTTATTACAG	1560
Qy	1468	GGATACCCCGTGAGTTAACCGCGGGCGCGCTTTGGCGTAATCATGGTCACTAGCTGTTTCC	1527
Db	1561	CTTGAGTATTTCTATAGTGTCACTTAATAGCTTGGCGTAATCATGGTCACTAGCTGTTTCC	1620
Qy	1528	TGTGTGAATTTGTTATCCGCTCAAAATTCACACAACATACAGACCGGGAAGCATAAAGTG	1587
Db	1621	TGTGTGAATTTGTTATCCGCTCAAAATTCACACAACATACAGACCGGGAAGCATAAAGTG	1680
Qy	1588	TAAAGCCTGGGTGCCTAATAGTGAGCTAACTACATTAATTTGCGTTGCGTCACTGCC	1647
Db	1681	TAAAGCCTGGGTGCCTAATAGTGAGCTAACTACATTAATTTGCGTTGCGTCACTGCC	1740
Qy	1648	CGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACCGCGGG	1707
Db	1741	CGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACCGCGGG	1800
Qy	1708	GAGAGCGGTTTGGGTTATGGGCGCTCTTCCGCTTCTCGCTCATCTGACTCGCTGGGCTC	1767
Db	1801	GAGAGCGGTTTGGGTTATGGGCGCTCTTCCGCTTCTCGCTCATCTGACTCGCTGGGCTC	1860
Qy	1768	GCTCGTTTCGCTCGCGGAGCGGTATCAGCTCACTCAAGCGGCTAATACGTTATTCCAC	1827
Db	1861	GGTCTGTTTCGCTCGCGGAGCGGTATCAGCTCACTCAAGCGGCTAATACGTTATTCCAC	1920
Qy	1828	AGAAATCAGGGGATAACCCAGGAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA	1887
Db	1921	AGAAATCAGGGGATAACCCAGGAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA	1980
Qy	1888	CCGTAAAAAGCCCGCTGTGTCGGTTTTTTCATATAGGCTCCGCCCCCTGACGAGCATCA	1947
Db	1981	CCGTAAAAAGCCCGCTGTGTCGGTTTTTTCATATAGGCTCCGCCCCCTGACGAGCATCA	2040
Qy	1948	CHAAAATCGAGCTCAAGTCAGAGTCGGCGAAACCCGACAGGACTATAAGATACACAGGC	2007
Db	2041	CAAAAATCGAGCTCAAGTCAGAGTCGGCGAAACCCGACAGGACTATAAGATACACAGGC	2100
Qy	2008	GTTTCCCCCTGGAAAGCTCCCTCGTGGCGCTCTCTCTGTTCCGACCTCGCGCTTACCCGATCA	2067

[illegible]

Db	3181	AAAAGCGTTAGCTCCTTCGGTCTCGATCGTTGTGAGAGTAAGTTGGCCGCACTGTT	3240
QY	3148	ATCAGTATGTTATGCGAGCACTCATAAATCTCTTACTGTGATGCCATCCGATCGATG	3207
Db	3241	ATCAGTATGTTATGCGAGCACTCATAAATCTCTTACTGTGATGCCATCCGATCGATG	3300
QY	3208	CTTTCTGAGTCTGTTGAGTACTCAACCAAGTCAATCTGAGATAGTGTATGCGGCGAC	3267
Db	3301	CTTTCTGAGTCTGTTGAGTACTCAACCAAGTCAATCTGAGATAGTGTATGCGGCGAC	3360
QY	3268	GAGTTGCTCTTGGCCGCGCTCAATACGGGATATACCGGCCACATAGCAGAACTTTTAA	3327
Db	3361	GAGTTGCTCTTGGCCGCGCTCAATACGGGATATAGTGTATGACATAGCAGAACTTTA	3420
QY	3328	AGTGTCTATGTTGAAAACGTTCTTTCGGGGCGAAAACCTCTCAAGATCTTACCGCTGT	3387
Db	3421	AGTGTCTATGTTGAAAACGTTCTTTCGGGGCGAAAACCTCTCAAGATCTTACCGCTGT	3480
QY	3388	GAGATCCAGTCTGATTAACCACTCGTGCACCCCACTGATCTTCAGCATCTTTTACTTT	3447
Db	3481	GAGATCCAGTCTGATTAACCACTCGTGCACCCCACTGATCTTCAGCATCTTTTACTTT	3540
QY	3448	CACCAAGCTTCTGGGTGAGCAAAAACAGAGGCAAAAATGCCGCAAAAAGGGAATAAG	3507
Db	3541	CACCAAGCTTCTGGGTGAGCAAAAACAGAGGCAAAAATGCCGCAAAAAGGGAATAAG	3600
QY	3508	GCGACACGGAATGTTGAATCTATCATCTCTTCTTTTCAATATTTAGAGCAATTA	3567
Db	3601	GCGACACGGAATGTTGAATCTATCATCTCTTCTTTTCAATATTTAGAGCAATTA	3660
QY	3568	TCAGGGTATGTTCTCATGAGCGGATACATATTTGATGATGTTTGAATAATAACAAAT	3627
Db	3661	TCAGGGTATGTTCTCATGAGCGGATACATATTTGATGATGTTTGAATAATAACAAAT	3720
QY	3628	AGGGGTTCCGCGCACATTTTCCCGAAAAGTGCCAC	3662
Db	3721	AGGGGTTCCGCGCACATTTTCCCGAAAAGTGCCAC	3755
RESULT 11			
AF179627			
LOCUS	AF179627	5540 bp	DNA linear SYN 22-NOV-1999
DEFINITION	Expression vector pCENEX645, complete sequence.		
ACCESSION	AF179627		
VERSION	AF179627.1	GI:6457303	
KEYWORDS	Expression vector pCENEX645		
SOURCE	Expression vector pCENEX645		
ORGANISM	Other sequences; artificial sequences; vectors.		
REFERENCE	1 (bases 1 to 5540)		
AUTHORS	Eugster,H.P., Bartsch,S., Murgler,F.E. and Sengstag,C.		
TITLE	Functional co-expression of human oxidoreductase and cytochrome P450 1A1 in Saccharomyces cerevisiae results in increased EROD activity		
JOURNAL	Biochem. Biophys. Res. Commun. 185 (2), 641-647 (1992)		
MEDLINE	92304288		
PUBMED	1610357		
REFERENCE	2 (bases 1 to 5540)		
AUTHORS	Sengstag,C. and Paladino,G.		
TITLE	A collection of transformed Saccharomyces cerevisiae strains as an unlimited source of human enzymes		
JOURNAL	BioWorld 97 (6), 2-7 (1997)		
REFERENCE	3 (bases 1 to 5540)		
AUTHORS	Sengstag,C.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-AUG-1999) Genetics, Institute of Toxicology, ETH Zurich, Schorenstrasse 16, Schwerzenbach 8603, Switzerland		
FEATURES	Location/Qualifiers		
source	1..5540		
	/organism="Expression vector pCENEX645"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:106613"		
	/note="Yeast expression vector pCENEX645 (alias pCS645)"		

		containing truncated GAPDH promoter for constitutive expression of cDNAs in yeast"	
		417..1220	
		/gene="URA3"	
		417..1220	
		/gene="URA3"	
		/note="yeast selection marker"	
		/codon_start=1	
		/product="URA3"	
		/db_xref="GI:6457304"	
		/translation="MSKATYKERAATHPSVAAKLFNIMHEKQTNLCASLDVRRTKEL	
		LELVALGPKICLLATHVDILTDFSMEGTVKPKALKSAKNFLLEDKFEADIGNTVK	
		LQYSGVTRIAEWADITNAHGVVGGIVSGLKQAABEVTKPRGLLMLAELCKGSLIS	
		TGEYTKGTVDIAKSDKFVIGTIAQRDMGDRGDEYDLWLTMTGVLGDLDDIGDALGQQYR	
		TVDDVVSFGSDIIIVGRFLAFKGRDAKVEGERYRKAGWEAYLRRCQQN"	
		2310..2508	
		/note="truncated GAPDH promoter"	
		complement(3958..4818)	
		/gene="Amp"	
		complement(3958..4818)	
		/gene="Amp"	
		/note="ampicillin resistance; Escherichia coli selection marker"	
		/codon_start=1	
		/product="Amp"	
		/protein_id="AAF09467.1"	
		/db_xref="GI:6457305"	
		/translation="MSIOHFRVALIPFFAAFLPVFAHPDLLVFKVDCDQLGARVGY	
		IELDLSGKILESPEPSPMMSTFKVLLCGAVLPRIDAGOGQLGRRTHYSQNDLVE	
		YSPVTEKHLTDGMTVRELCSAALTMSDNTAANLLTTIGGPKELTFLNMDGHDVTRL	
		DRWPELNEALPNDERDITMPVAMATTLRLKLTGELLTLASRQQLIDWNEADKVAGPL	
		LRSALPAKWFADKSGAGERSGRIIAALGPDGKPSRIVVIYTTGSGQATMDERNROIA	
		EIGASLIKHW"	
ORIGIN			
		Query Match 69.9%; Score 2558.8; DB 12; Length 5540;	
		Best Local Similarity 82.8%; Pred. No. 0;	
		Matches 3035; Conservative 0; Mismatches 557; Indels 72; Gaps 7;	
QY	3	AAATTTGTAAGCGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCAT	62
Db	1351	AAATTTGTAAGCGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCAT	1410
QY	63	TTTTTAACCAATAGCCGAAATCGGCAAAATCCCTTATAAATCAAAGAATAGACCGAGA	122
Db	1411	TTTTTAACCAATAGCCGAAATCGGCAAAATCCCTTATAAATCAAAGAATAGACCGAGA	1470
QY	123	TAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAGACGTGGACTCCA	182
Db	1471	TAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAGACGTGGACTCCA	1530
QY	183	ACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCT	242
Db	1531	ACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCT	1590
QY	243	AATCAAGTTTTTGGGGTGCAGGTCCGCTAAAGCACTAAATCGGAACCTCTAAAGGAGCC	302
Db	1591	AATCAAGTTTTTGGGGTGCAGGTCCGCTAAAGCACTAAATCGGAACCTCTAAAGGAGCC	1650
QY	303	CCGATTTAGACTTTGACGGGAAACCGCGCAAGTGGCGGAGAAAGGAGGAAAG	362
Db	1651	CCGATTTAGACTTTGACGGGAAACCGCGCAAGTGGCGGAGAAAGGAGGAAAG	1710
QY	363	CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAAACACCA	422
Db	1711	CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAAACACCA	1770
QY	423	CACCCGCGCGCTTAATGCGCGCTACAGGGCGCGTC---CCATTGCGCATTCAGGCTGC	479
Db	1771	CACCCGCGCGCTTAATGCGCGCTACAGGGCGCGTC---CCATTGCGCATTCAGGCTGC	1830
QY	480	GCAACTGTTGGGAAGGGCGATCGGTGCGGGCGCTCTTCGCTATTACGCCAGCTGGCGAAG	539

1831 GCACCTGTTGGAAAGGGCGATCGGTGGGCGCTCTTCGTATTACCGCAGCTGGCGAAAG 1890
540 GGGGATGCTGCAAGGCGATTAAAGTTGGTTAAAGCCAGGGTTTCCAGTCAACAGCTT 599
1891 GGGGATGCTGCAAGGCGATTAAAGTTGGTTAAAGCCAGGGTTTCCAGTCAACAGCTT 1950
600 GTAAACGAGCGCCAGTGAAGCGCGCTCGTTCAATTCACGTTTTTGAACCCGTGGAGGAG 659
1951 GTAAACGAGCGCCAGTGA-----ATTGTAATACGACTCACTATAGGCGGAATTGG 2001
660 GGCAGACTCGCGGTGCAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACA 719
2002 AGCTCCACCGCGGTG-----GCGCGCGCTCTAGTCGACGCTCTCC 2042
720 CGCTCGAGAACACGAGCTAGATTAAACCTAGAAAGATAATCATATTGTGAGCTAGCTTA 779
2043 TTATCGCACTCTCGATTAGGAAGCAGCCAGTAGTAGTTGAGGCCGTTGAGCACCGCC 2102
780 AAGATAATCATCGGTAAATAATTCACGCATGGGATCTGTAATACGACTCACTATAGGCGAA 839
2103 GCCGCAAGGAATGGTGCAATGAAGAGATGGCGCCCAACAGTCCCGGCCACGGGCGCT 2162
840 TTGGGTACCGGCCCCCGCTCGAGTTCGACGGTATCGATAAGCTTTGATATCGAATTCCTG 899
2163 GCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCGAAGTGGCGAGCCCGATCTTCC 2222
900 CAGCCCGGGGATCCACTAGTTCTAGAGCGGCGCCACCGCGGTGGAGCTCAGCTTTTG 959
2223 CCATCGGTGATCGCGCGATATAGGCGCGAGCAACCGCACCTGTGGCGCGGTGATGCG 2282
960 TTCCCTTTAGTGAGGGTTAATTAGATCCCATCGCTCAATTTTACGACGACTATCTTTCTA 1019
2283 GCCACGATGCGTCCGCGTAGAGGATCTCCCTGAAATAAGGTTGAAACAGTTCCCTG 2342
1020 GGGTTAATCTAGCTGCATCAGGATCATATCGTGGGTCTTTTCCGGTCAAGTCAATCGC 1079
2343 AAATTTATTC-----CCCTACTTGACTAATAAGTATATA 2376
1080 CCAAGCTGGCGTATCTGGGATTCGGGAGGAGGAGCCCGTCTTTTCCCGGAGGTT 1139
2377 AGACGGTAGGTATGATTGTAAATCTGTA-----AATCTATTCTTAAACTTCTTAAAT 2431
1140 GAAGCGGATGAAGAGTTTCCGAGGATGACTGCTGCTGATGACGTTGAGCGGAA 1199
2432 CTACTTTATAGTTAGTCTTTTITTAGTT-----TTAAACACCAAGAACTTAGT 2482
1200 CGCAGTTTACCATGATTCGGGAAGGTGGCCATGCAAGCTTTTAAACGGTGAAC 1259
2483 TTCGAATAAACACACATAAATAAGAAATTCATGCTTTTGCMAAGCTTTCCTTTTCCCTTTG 2542
1260 TTCGTTCAAGCCACCTGGGATACAGTTGCTCGCGCTTTTCCGACACAGTTCCGGATG 1319
2543 GCTGTTTTTGCAGCAAAATATCTGCATCTTAGCTCTCGAGGGAACACTACTCATTAC 2602
1320 GTCAGCCGGAAGCGATCAGCAACCGAAACATAACCGGACAGCGGAACTGCGCTGCC 1379
2603 AACGCCAGTCTATTGACACAATAGTTTGTATACTAATAATATGGAACATAATAGC 2662
1380 GGTGTGCAATTATGACAGCGGTGGCGCTGGGATATTAAGTCAGCGAGGACGGGTAT 1439
2663 AATACCCAAATTTTATCTAAATTTTCCGAAAGATTAAATTAATCAGCTTATCGATACCGT 2722
1440 CTTGCTGATCCCGCAGAAATGGACATGGATACCCCGTAGTTACCGG-GCGGGCGGC 1498
2723 CGAGGGGGGCGCGCGCGCCCGCTTTGTTCCCTTTAGTGGGGTTAATCCGAGC 2782
1499 TTGGCGTAATCATGCTCATAGCTGTTTCTGTTGAAATTTGTTATCCGCTCAACAATCCA 1558
2783 TTGGCGTAATCATGCTCATAGCTGTTTCTGTTGAAATTTGTTATCCGCTCAACAATCCA 2842
1559 CACAAATACGAGCGGAGCATATAAGTGAAGCTTGGGTCCTTAATGATGAGCTAA 1618

2843 CACAAATACGAGCGGAGCATATAAGTGTAAAGCCTGGGGTGCCTAATGATGAGGTAA 2902
1619 CTCACATTAATTTGCGTTGCGTCACTGCCCGCTTTTCCAGTCGGGAAACCTGTCGTGCCAG 1678
2903 CTCACATTAATTTGCGTTGCGTCACTGCCCGCTTTTCCAGTCGGGAAACCTGTCGTGCCAG 2962
1679 CTGCAATTAATGAAATCGGCAACCGCGGGGAGCGGTTTTCGTTATTTGGGCGCTCTTCC 1738
2963 CTGCAATTAATGAAATCGGCAACCGCGGGGAGCGGTTTTCGTTATTTGGGCGCTCTTCC 3022
1739 GCTTCCTCGCTCACTGACTCGCTGCGTCTGCTGCTGCTGCGGCGAGCGGTATCAGCT 1798
3023 GCTTCCTCGCTCACTGACTCGCTGCGTCTGCTGCTGCTGCGGCGAGCGGTATCAGCT 3082
1799 CACTCAAAAGGGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGAGGAAGAACATG 1858
3083 CACTCAAAAGGGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGAGGAAGAACATG 3142
1859 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGTTGCTGGCGCTTTTTC 1918
3143 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGCTTTTTC 3202
1919 CATAGGCTCGCGCCCGCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGA 1978
3203 CATAGGCTCGCGCCCGCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGA 3262
1979 AACCCGACAGGACTATAAGATACAGGCTTTTCCCGCTGGAAGCTCCCTCGTGGCTCT 2038
3263 AACCCGACAGGACTATAAGATACAGGCTTTTCCCGCTGGAAGCTCCCTCGTGGCTCT 3322
2039 CCGTTTCGACCCCTCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAAGCGTG 2098
3323 CCGTTTCGACCCCTCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAAGCGTG 3382
2099 GCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAG 2158
3383 GCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAG 3442
2159 CTGGGCTGTGTGACGAACCCCGCTTCAAGCCGACCGCTGCTTATCCCGTAACTAT 2218
3443 CTGGGCTGTGTGACGAACCCCGCTTCAAGCCGACCGCTGCTTATCCCGTAACTAT 3502
2219 CGTCTTGAGTCAACCCCGTAAACACGACTTATCGCACTGCGCAGCAGCCACTGTAAC 2278
3503 CGTCTTGAGTCAACCCCGTAAACACGACTTATCGCACTGCGCAGCAGCCACTGTAAC 3562
2279 AGGATTACGAGCGAGGTATGTAGCGGTGCTCAGAGTCTTCAAGTGTGGCTTAAC 2338
3563 AGGATTACGAGCGAGGTATGTAGCGGTGCTCAGAGTCTTCAAGTGTGGCTTAAC 3622
2339 TACGGCTACACTAGAAAGGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTC 2398
3623 TACGGCTACACTAGAAAGGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTC 3682
2399 GGAATAAGAGTTGTTAGCTTTGATCCGCAACAAACACCGCTGGTAGCGGTGTTTT 2458
3683 GGAATAAGAGTTGTTAGCTTTGATCCGCAACAAACACCGCTGGTAGCGGTGTTTT 3742
2459 TTTGTTTGAAGCAGCAGATTACCGCGAGAAAAAGGATCTCAAGAGATCCTTTGATC 2518
3743 TTTGTTTGAAGCAGCAGATTACCGCGAGAAAAAGGATCTCAAGAGATCCTTTGATC 3802
2519 TTTTCTACGGGTCTGAGCGCTCAGTGGAAACGAAACTCAGCTTAAGGGATTTTGGTCATG 2578
3803 TTTTCTACGGGTCTGAGCGCTCAGTGGAAACGAAACTCAGCTTAAGGGATTTTGGTCATG 3862
2579 AGATTATCAAAAGGATCTTCACTCCTAGATCTTTTAAATTTAAATTAAGTATTAATCA 2638
3863 AGATTATCAAAAGGATCTTCACTCCTAGATCTTTTAAATTTAAATTAAGTATTAATCA 3922
2639 ATCTAAGTATATAGTAACTTTGCTGACAGTTACCAATGCTTAATCAGTGGGCA 2698
3923 ATCTAAGTATATATAGTAACTTTGCTGACAGTTACCAATGCTTAATCAGTGGGCA 3982

QY	2699	CTATCTCAGCGATCTGTCTATTTCTTCATTCATATAGTTGCTCTGACTCCCGTCTGRTAG	2758
Db	3983	CTATCTCAGCGATCTGTCTATTTCTTCATTCATATAGTTGCTCTGACTCCCGTCTGRTAG	4042
QY	2759	ATAAATACCATACGGGAGGCTTACCATTCTGGCCCGAGTCTGCAATGATACCGCGAGAC	2818
Db	4043	ATAAATACCATACGGGAGGCTTACCATTCTGGCCCGAGTCTGCAATGATACCGCGAGAC	4102
QY	2819	CCACGCTACCGGCTCCAGATTATCAGCAATAAACACGACCGCGGAGGCGCGAGCGC	2878
Db	4103	CCACGCTACCGGCTCCAGATTATCAGCAATAAACACGACCGCGGAGGCGCGAGCGC	4162
QY	2879	AGAAGTGTCTCGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGGAAGCT	2938
Db	4163	AGAAGTGTCTCGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGGAAGCT	4222
QY	2939	AGAGTAAGTAGTTCCGCACTTAATAGTTGCGCAAGTGTGTCGCAATTCGTACAGGCATC	2998
Db	4223	AGAGTAAGTAGTTCCGCACTTAATAGTTGCGCAAGTGTGTCGCAATTCGTACAGGCATC	4282
QY	2999	GTGGTGTACGCTCGTCTGTTGGTATGGCTTCATTCAGCTCCGGTTCCTCAACGATCAAGG	3058
Db	4283	GTGGTGTACGCTCGTCTGTTGGTATGGCTTCATTCAGCTCCGGTTCCTCAACGATCAAGG	4342
QY	3059	CGAGTTACATGATCCCGCATGTTGTCGCAAAAAGCGGTTAGTCTCTCGGTCCTCGGATC	3118
Db	4343	CGAGTTACATGATCCCGCATGTTGTCGCAAAAAGCGGTTAGTCTCTCGGTCCTCGGATC	4402
QY	3119	GTGTGCAAGTAGTTGCGCGCAGTGTATCACTCATGTTATGCGCAGCTGATAT	3178
Db	4403	GTGTGCAAGTAGTTGCGCGCAGTGTATCACTCATGTTATGCGCAGCTGATAT	4462
QY	3179	TCTTCTACTGTATGTCATCCGTAAGTCTTTCTGTGACTGGTGGTACTCAACCAAG	3238
Db	4463	TCTTCTACTGTATGTCATCCGTAAGTCTTTCTGTGACTGGTGGTACTCAACCAAG	4522
QY	3239	TCATTCTGAGATAGTGTATGCGGACCGAGTGTCTTTCGCGCGCTCAATACGGGAT	3298
Db	4523	TCATTCTGAGATAGTGTATGCGGACCGAGTGTCTTTCGCGCGCTCAATACGGGAT	4582
QY	3299	ATATCCGCCACATAGCAGAACTTTAAAGTCTCATCTATTTGGAAGCTTCTTCGGG	3358
Db	4583	ATATCCGCCACATAGCAGAACTTTAAAGTCTCATCTATTTGGAAGCTTCTTCGGG	4642
QY	3359	CGAAACTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTCGATGTAACCCACTCGTGCA	3418
Db	4643	CGAAACTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTCGATGTAACCCACTCGTGCA	4702
QY	3419	CCCAACTGATCTTCAGCATCTTTTACTTTTCCAGCGCTTCTGGGTGACGCAAAAACAGGA	3478
Db	4703	CCCAACTGATCTTCAGCATCTTTTACTTTTCCAGCGCTTCTGGGTGACGCAAAAACAGGA	4762
QY	3479	AGGCAAAATGCGCAAAAAGGAATAAGGCGGACACGGAATGTTGATATCTCATCTC	3538
Db	4763	AGGCAAAATGCGCAAAAAGGAATAAGGCGGACACGGAATGTTGATATCTCATCTC	4822
QY	3539	TTCTCTTTTCAATATTATTGAAGCATTTTATCAGGGTTATTTGCTCATGAGCGGATACATA	3598
Db	4823	TTCTCTTTTCAATATTATTGAAGCATTTTATCAGGGTTATTTGCTCATGAGCGGATACATA	4882
QY	3599	TTTGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGCGACATTTTCCCGGAAAAAGTG	3658
Db	4883	TTTGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGCGACATTTTCCCGGAAAAAGTG	4942
QY	3659	CCAC 3662	
Db	4943	CCAC 4946	
RESULT 12			
ASPGREEN2/c			
LOCUS			
3633 bp DNA linear SYN 15-MAY-1998			

DEFINITION	Artificial sequences, plasmid vector pGreen-2.
ACCESSION	Y09374
VERSION	Y09374.1 GI:1684627
KEYWORDS	blatEM gene; lacZ-PhoC fusion gene; multiple cloning site; T3 promoter; T7 promoter.
SOURCE	synthetic construct
ORGANISM	other sequences; artificial sequences.
REFERENCE	1
AUTHORS	Thaller, M.C., Berlutti, F., Schippa, S., Selan, L. and Rossolini, G.M.
TITLE	Bacterial acid phosphatase gene fusions useful as targets for cloning-dependent insertional inactivation
JOURNAL	Biotechnol. Prog. 14 (2), 241-247 (1998)
MEDLINE	98215688
PUBMED	9548775
REFERENCE	2 (bases 1 to 3633)
AUTHORS	Rossolini, G.M.
TITLE	Direct Submision
JOURNAL	Submitted (11-NOV-1996) G.M. Rossolini, Dip. Biologia
COMMENT	Moleculare-Sez., Microbiologia,, Univ. di Siena, via Laterina N.8, I- 53100 Siena, ITALY
FEATURES	Related sequences: X6444 and X52328.
source	1..3633
	Location/Qualifiers
	/organism="synthetic construct"
	/mol_type="other DNA"
	/db_xref="taxon:32630"
	/plasmid="pGreen-2"
	129..989
gene	/gene="blatEM"
CDS	129..989
	/gene="blatEM"
	/codon_start=1
	/transl_table=11
	/protein_id="CAA70549.1"
	/db_xref="GI:1684628"
	/translation="MSIQHFRVALLPFFAFCFLPVFAHPETLVKVDADQLGARVGY
	IEDLNSGKILDESFRPEFPMWTFKLLCGAVLSRIDAGQQLGRRIRHYSNDLVE
	YSPVTEKHLDTGMTRVLCSSAATMDSNTAANLLLTIGGPKELTAFLHMGDVTRL
	DRWPELNEATPNDRDRTMPVAMATTLRKLLTGELLTLASRQLITDMNEADKVAQPL
	LRSLALPAGWFIADKSGAGERSGRIIAALGPDGKPSRIIVIVTTGSGATMDENRQIA
	ETGASLIKHM"
primer_bind	2147..2162
	/note="PUC M13/Reverse Sequencing primer recognition site"
gene	2154..3104
	/gene="lacZ-PhoC"
CDS	2154..3104
	/gene="lacZ-PhoC"
	/codon_start=1
	/transl_table=11
	/protein_id="CAA70550.1"
	/db_xref="GI:1684629"
	/translation="MTMITPSAQLTLTKGNKSWSTAVAAALELVDPGPCRNSISLS
	IPSTRGPGVPNSPYSESYARS LAVLQRRDWPNGVTHGAIPAGNDATTPDLYYL
	KNEAIDSLKLPPEVGSIQFLNDQMYERMLRTERGKQAQADLAAGVAT
	AFSGAFPIPTKDPKLPYKLTNMIEDAGDILATRSKEHYMRIIPFAYGTCTNK
	DKDLSTGSPSGHTSIQWATLVLAEPNPAQDAILERGYQLGSRVICGVHMQSD
	VDAARI VGSAAVATLHSDPAFOAQLAKAKQBEFAQKSQDL"
promoter	2179..2198
	/gene="lacZ-PhoC"
	/note="T3 promoter"
misc_feature	2214..2316
	/gene="lacZ-PhoC"
	/note="multiple cloning site"
promoter	2325..2344
	/gene="lacZ-PhoC"
primer_bind	2354..2370
	/note="T7 promoter"
	/gene="lacZ-PhoC"
primer_bind	2374..2390
	/note="M13-20 sequencing primer recognition site"
	/gene="lacZ-PhoC"
	/note="M13-40 sequencing primer recognition site"

ORIGIN			
Query Match 69.7%; Score 2550.6; DB 12; Length 3633;			
Best Local Similarity 82.3%; Pred. No. 0;			
Matches 3020; Conservative 0; Mismatches 604; Indels 47; Gaps 6;			
QY	1	CTAAATGTTAGCGGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC	60
DB	3633	CTAAATGTTAGCGGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC	3574
QY	61	ATTTTAAACCAATAGCGGAAATCGGAAATCCCTTATAAATCAAAAGATAGACGA	120
DB	3573	ATTTTAAACCAATAGCGGAAATCGGAAATCCCTTATAAATCAAAAGATAGACGA	3514
QY	121	GATAGGTTGAGTGTTCAGTTTGGAAACAAGAGTCCACTATTAAAGAACTGGACTC	180
DB	3513	GATAGGTTGAGTGTTCAGTTTGGAAACAAGAGTCCACTATTAAAGAACTGGACTC	3454
QY	181	CAACGTCAAAGGGCGAAACCGTCTATCAGGGCGATGGCCACTACGTGAACCATCACC	240
DB	3453	CAACGTCAAAGGGCGAAACCGTCTATCAGGGCGATGGCCACTACGTGAACCATCACC	3394
QY	241	CTAATCAAGTTTTTGGGGTCAGGTGCGGTAAGCACTAAATCGGAACCTTAAAGGAG	300
DB	3393	CTAATCAAGTTTTTGGGGTCAGGTGCGGTAAGCACTAAATCGGAACCTTAAAGGAG	3334
QY	301	CCCCGATTTAGAGCTTGACGGGGAAGCGCGCAACGTGGCGAGAAAGGAAGGAAGAA	360
DB	3333	CCCCGATTTAGAGCTTGACGGGGAAGCGCGCAACGTGGCGAGAAAGGAAGGAAGAA	3274
QY	361	AGCGAAAGAGCGGGCGCTGAGCGCTGCAAGTGTAGCGGTACGCTCGCGTTAACCAAC	420
DB	3273	AGCGAAAGAGCGGGCGCTGAGCGCTGCAAGTGTAGCGGTACGCTCGCGTTAACCAAC	3214
QY	421	CACACCGCGCGCTTAATGCGCGGTCACAGGCGGTCGCCATTGCGCAATTCAGGCTGG	480
DB	3213	CACACCGCGCGCTTAATGCGCGGTCACAGGCGGTCGCCATTGCGCAATTCAGGCTGG	3154
QY	481	CAACTGTTGGGAAGGCGCATCGGTCGGGCTCTTCGCTATTACGCCAGCTGGCGAAAGG	540
DB	3153	CAACTGTTGGGAAGGCGCATCGGTCGGGCTCTTCGCTATTACGCCAGCTGGCGAAAGG	3094
QY	541	GGGATGTGTCGAAGCGGATTAAGTTGGGTAAACGCCAGGCTTTCCAGTACGACGTTG	600
DB	3093	TCTGTGATTTTGTGCAAAATCCTGTTGGCTTCGCTAACTGCGCTGAAATGCGGAT	3034
QY	601	TAAACGACGCGCAGTACGCGGCTCGTTCAATTCAGCTTTTGAACCGGTGGAGACGG	660
DB	3033	CGGAATGTAATGTCGCGACAGCGGCTGAACCGACAATCGCGCGGCATCCACATCACTCT	2974
QY	661	GCAGACTCGCGTGCAAAATGTGTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC	720
DB	2973	GCCAGTGATAGCGGCAAAATCACCC-----GGCTCTGTCGAGCTGATTAACCCC	2926
QY	721	GCTGCAGAACCGCAGCTAGATTAAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAA	780
DB	2925	GTTCCAGAATCGCATCTGATTTGCGGGTTCACCTTCGCGCAGCACTAATGCGGTGGCC	2866
QY	781	AGATAATCATGCGTAAATTTAGCATGGATCTGTATAACACTCACTATAGGCGGAAT	840
DB	2865	AGCCGATAGCATATGACCTCGCGGTAAAGTCCGTTGGTGGAGAGTTTTTCTGATCTT	2806
QY	841	TGGGTACCGGGCCCCCTCGAGGTCGAGGTATCGATAAGCTTGATATCGAATTCCTGTC	900
DB	2805	TGGTATTACAGGTTTCTGTGCGGTAACCGAAACGAGCGGATGCGCATGTAATGTTCTT	2746
QY	901	AGCCCGGGGATCCACTAGTTCTTAGAGCGCGCGCCACCGCGGTGGAGCTTCCAGCTTTTGT	960
DB	2745	TGCGGAGCGGTTGCAAGGTCACCGGCATCCTCAATCATATTGTTGACAGTTTATACA	2686
QY	961	TCCCTTTAGTAGGGTTAATTAGATCCCATGCGTCAATTTTACGAGCATATCTTTCTAG	1020
DB	2685	GCTCCGGAGAGTCTTTTTCGGTTATCGGATAGCCGAATG--CCCCGAAATGCGGTTGC	2628

QY	1021	GTTAATCTAGCTGCATCAGGATCATATCTGTCGGGTCTTTTTTTCGGCTCAGTCATCGCC	1080
DB	2627	CACCCCCCTCGGCCAGGTGAGCATCTGCTGTGCTGTTTTCGCGCTGCTATGCG	2568
QY	1081	CAAGTCGGCGGTATCTGGGATCGGGAGGAAGACCGGTCCTTTTCCCGGAGGTTG	1140
DB	2567	CAGCATACGGGCTTTCTCATATCATTTGCCCTGATCATTTAAAACTGAATCTGCCAGCTTC	2508
QY	1141	AAGCGG-CATGGAAGAGTTTCCGAGGATGACTCTGCTGCATTGACGTTTGAGCGAAA	1199
DB	2507	CGCGGTCGGGTAAACAGTTTCAGGCTGTGATAGCTGTTCAITTTTTCAGATAATATAA	2448
QY	1200	CGCAGCTTTTACCATGATGATTCGGGAAGGTGTGGCCATCATCGACGCTTTTAAACGGTGAAC	1259
DB	2447	ATCCGGCTTGGTGGGATCGTTGCTCGCGGATCGTCCATGGTAAACCCAGGGTT	2388
QY	1260	TTGCTTACGGCACTCGGGATACAGTTGTCGCGGCTTTTCCGGACACAGTTTCGGATG	1319
DB	2387	TTCCAGTACGACGCTTTGTAACACGACGCGCAGTACGCGCGGTAAATACGACTACTATA	2328
QY	1320	GTACGCCCCGAGCGCATCAGCAACCCGAAACAATACCGCGACAGCGGAACTGCCGTGCC	1379
DB	2327	G-----GGCGAATTTGGGTACCGGGCCCCCTCGAGGTCGACGCTATCGATAAGCT	2277
QY	1380	GGTGTGCAGATTAATGACAGCGGTGCGGCTGGGATATTACGTACGCGAGGACCGGTAT	1439
DB	2276	TGATATCGAATTCCTGCAGCGCGGGGATCCACTAGTTTCTAGAGCGCGCCACCGCGT	2217
QY	1440	CCTGCTGGATGCCGCAAAATGGACATGGATACCCCGTGTAGTTACCCGCGGCGCGCT	1499
DB	2216	GGAGCT-----CCAGCTTTTGTTCCTTTTAGTGAGGGTTAAATTTGCGCGCT	2172
QY	1500	TGGCTTAATCATGTGCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTCAC	1559
DB	2171	TGGCTTAATCATGTGCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTCAC	2112
QY	1560	ACAAATACGACCGCGGAAGCATAAAGTGTAAAGCCTCGGGTGCCTTAATGAGTGAGCTAAC	1619
DB	2111	ACAAATACGACCGCGGAAGCATAAAGTGTAAAGCCTCGGGTGCCTTAATGAGTGAGCTAAC	2052
QY	1620	TCATTAATTTGCTGTCCTCAGTCCGCGCTTTTCCAGTCCGGAACCTGTCTGCCAG-	1678
DB	2051	TCATTAATTTGCTGTCCTCAGTCCGCGCTTTTCCAGTCCGGAACCTGTCTGCCAGC	1992
QY	1679	-----CTGCATTAATGAATCGGCAACCGCGGGGAGAGCGGTTTTCGCTATTGGGCG	1731
DB	1991	TGCATGCTGCAATTAATGAATCGGCAACCGCGGGGAGAGCGGTTTTCGCTATTGGGCG	1932
QY	1732	CTCTTCGCTTCTCTGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1791
DB	1931	CTCTTCGCTTCTCTGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1872
QY	1792	ATCAGCTCATCAAAAGGCGGTAAATAGGTTATCCACAGATCAGGGGATAACGCGAGAA	1851
DB	1871	ATCAGCTCACTCAAAAGGCGGTAAATAGGTTATCCACAGATCAGGGGATAACGCGAGAA	1812
QY	1852	GAAATGTGACCAAAAGCGCAGCAAAAGCGAGGAACCGTAAAGGCGCGCTGCTGCGC	1911
DB	1811	GAAATGTGACCAAAAGCGCAGCAAAAGCGAGGAACCGTAAAGGCGCGCTGCTGCGC	1752
QY	1912	GTTCATGAGGCTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1971
DB	1751	GTTCATGAGGCTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1692
QY	1972	GTGGGAAACCGCAGAGCATATAAGATACAGGCGTTTCCCTTGGAGCTCCCTCGT	2031
DB	1691	GTGGGAAACCGCAGAGCATATAAGATACAGGCGTTTCCCTTGGAGCTCCCTCGT	1632
QY	2032	GCGCTCTCTGTTTCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGG	2091
DB	1631	GCGCTCTCTGTTTCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGG	1572


```
QY 2092 AAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTGAGTGGTCTTCCG 2151
Db 1571 AAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTGAGTGGTCTTCCG 1512
QY 2152 CTCGAAGCTGGCTGTGTGTCAGCAACCCCGCTTACAGCCCGCAGCGCTGCGCTTATCCGG 2211
Db 1511 CTCGAAGCTGGCTGTGTGTCAGCAACCCCGCTTACAGCCCGCAGCGCTGCGCTTATCCGG 1452
QY 2212 TAACTATCGTTTGTAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGCGCAGCAGCCAC 2271
Db 1451 TAACTATCGTTTGTAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGCGCAGCAGCCAC 1392
QY 2272 TGGTAACAGGATAGCAGACGAGTATGTAGCGGTGCTACAGATTCTTGAAGTGGTG 2331
Db 1391 TGGTAACAGGATAGCAGACGAGTATGTAGCGGTGCTACAGATTCTTGAAGTGGTG 1332
QY 2332 GCCTAACTACGGCTACACTAGAGGACAGTATTTGGTATCTCGCTCTCTCAAGCCAGT 2391
Db 1331 GCCTAACTACGGCTACACTAGAGGACAGTATTTGGTATCTCGCTCTCTCAAGCCAGT 1272
QY 2392 TACCTTCGGAAAAAGAGTTGGTAGCTCTTGTATCCGGCAAAACAAACCCACCGCTGGTAGCGG 2451
Db 1271 TACCTTCGGAAAAAGAGTTGGTAGCTCTTGTATCCGGCAAAACAAACCCACCGCTGGTAGCGG 1212
QY 2452 TGGTTTTTTTGTGCAAGCAGCAGATTACGGCGCAGAAAAAAGGATCTCAAGAGATCC 2511
Db 1211 TGGTTTTTTTGTGCAAGCAGCAGATTACGGCGCAGAAAAAAGGATCTCAAGAGATCC 1152
QY 2512 TTTTGATCTTTTCTACGGGCTCTGAGCTCAGTGGAGCAAACTCAGCTTAAAGGATTTT 2571
Db 1151 TTTTGATCTTTTCTACGGGCTCTGAGCTCAGTGGAGCAAACTCAGCTTAAAGGATTTT 1092
QY 2572 GGTCTATGAGATTATCAAAAAGGATCTTCACTAGATCTCTTTTAAATTTAAATGAAGTTT 2631
Db 1091 GGTCTATGAGATTATCAAAAAGGATCTTCACTAGATCTCTTTTAAATTTAAATGAAGTTT 1032
QY 2632 TAAATCAATCTAAAGTATATATGATTAACCTTGGTCTGACAGTTACCAATGCTTAAATCAG 2691
Db 1031 TAAATCAATCTAAAGTATATATGATTAACCTTGGTCTGACAGTTACCAATGCTTAAATCAG 972
QY 2692 TGAGCACCCTACTCAGCGATCTGTCTATTTTCTGTTTCATCATAGTTGCTGACTCCCGT 2751
Db 971 TGAGCACCCTACTCAGCGATCTGTCTATTTCTGTTTCATCATAGTTGCTGACTCCCGT 912
QY 2752 CGTGTAGATAACTAGCATAACGGAGGGCTTACCATCTGCGCCAGTGTGCTCAATGATACC 2811
Db 911 CGTGTAGATAACTAGCATAACGGAGGGCTTACCATCTGCGCCAGTGTGCTCAATGATACC 852
QY 2812 GCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCGGAGGGC 2871
Db 851 GCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCGGAGGGC 792
QY 2872 CGAGCGCAGAGTGGTCTGCAACTTATTCGCGCTCCATCCAGTCTATTAAATGTTGCCG 2931
Db 791 CGAGCGCAGAGTGGTCTGCAACTTATTCGCGCTCCATCCAGTCTATTAAATGTTGCCG 732
QY 2932 GGAAGCTAGAGTAAGTAGTTCGCCAGTAAATAGTTTGGCAACAGTTGTTGCCATGCTAC 2991
Db 731 GGAAGCTAGAGTAAGTAGTTCGCCAGTAAATAGTTTGGCAACAGTTGTTGCCATGCTAC 672
QY 2992 AGGCATCGTGGTGTACCGCTCGTGGTGGTATGGCTTCATTCAGTCCGGTTCGCAAGC 3051
Db 671 AGGCATCGTGGTGTACCGCTCGTGGTGGTATGGCTTCATTCAGTCCGGTTCGCAAGC 612
QY 3052 ATCAGAGCGAGTTACATGATCCCGATGTTGTGCAAAAAGCGGTAGCTCTCGGTCC 3111
Db 611 ATCAGAGCGAGTTACATGATCCCGATGTTGTGCAAAAAGCGGTAGCTCTCGGTCC 552
QY 3112 TCCGATCGTGTGTGAGAAAGTAAAGTTGGCCGAGTGTATCACTCATGTTTATGGCAGCACT 3171
Db 551 TCCGATCGTGTGTGAGAAAGTAAAGTTGGCCGAGTGTATCACTCATGTTTATGGCAGCACT 492
QY 3172 GCATAATCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGATGTTGAGTACTC 3231
```

```
Db 491 GCATAATCTTCTTACTGTTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTC 432
QY 3232 AACCAAGTCATTCTCAGAAATAGTGTATGCGGCGACCGAGTGTCTCTTGGCCCGCGTCAAT 3291
Db 431 AACCAAGTCATTCTCAGAAATAGTGTATGCGGCGACCGAGTGTCTCTTGGCCCGCGTCAAT 372
QY 3292 ACGGATATAACCGCGCCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAAAGTTTC 3351
Db 371 ACGGATATAACCGCGCCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAAAGTTTC 312
QY 3352 TTCGGGGGGAACACTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTCAGTGTAAACCCAC 3411
Db 311 TTCGGGGGGAACACTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTCAGTGTAAACCCAC 252
QY 3411 TCGTCGACCCAACTGATCTTTCAGCATCTTTTACTTTTTCACAGCGTTTCTGGGTGAGCAAA 3471
Db 251 TCGTCGACCCAACTGATCTTTCAGCATCTTTTACTTTTTCACAGCGTTTCTGGGTGAGCAAA 192
QY 3472 AACAGAAAGGCAAAATGCGCAAAAAGGAAATAAGGGCGACACGGAATGTTGAATACT 3531
Db 191 AACAGAAAGGCAAAATGCGCAAAAAGGAAATAAGGGCGACACGGAATGTTGAATACT 132
QY 3532 CATACTCTCTCTTTTCAATATTTTGAAGCATTTATCAGGGTTATTGCTCATGAGCGG 3591
Db 131 CATACTCTCTCTTTTCAATATTTTGAAGCATTTATCAGGGTTATTGCTCATGAGCGG 72
QY 3592 ATACATATTTGAATGTATTAGAAAAATAAACAATAGGGTTTCCGCGCACATTTCCCGG 3651
Db 71 ATACATATTTGAATGTATTAGAAAAATAAACAATAGGGTTTCCGCGCACATTTCCCGG 12
QY 3652 AAAAGTGGCAC 3662
Db 11 AAAAGTGGCAC 1

RESULT 13
CQ768840/c
LOCUS
DEFINITION
ACCESSION
VERSION
CQ768840.1 GI:45112122
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
Xerion Pharmaceuticals GmbH (DB)
FEATURES
source
1..3918
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description of Artificial Sequence: plasmid"

ORIGIN

Query Match 68.8%; Score 2517.8; DB 6; Length 3918;
Best Local Similarity 81.8%; Pred. No. 0;
Matches 3052; Conservative 0; Mismatches 537; Indels 142; Gaps 7;

QY 3 AAATTTGAAGCGTAAATATTTTGTAAATTCGCGTTAAATTTTGTGTAAATCAGCTCAT 62
Db 3732 AAATTTGAAGCGTAAATATTTTGTAAATTCGCGTTAAATTTTGTGTAAATCAGCTCAT 3673
QY 63 TTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAAATCAAAAGAAATAGACCGAGA 122
Db 3672 TTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAAATCAAAAGAAATAGACCGAGA 3613
QY 123 TAGGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTATTATAAGAACGTTGAGTCTCA 182
```



```
QY 2272 TGTAAACAGGATTTAGCAGACGAGGATGATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTG 2331
Db 1463 TGTAAACAGGATTTAGCAGACGAGGATGATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTG 1404
QY 2332 GCCTAACTACGGCTACACTAGAGGACGATATTTGGTATCTGGCTCTGCTGAAGCCAGT 2391
Db 1403 GCCTAACTACGGCTACACTAGAGGACGATATTTGGTATCTGGCTCTGCTGAAGCCAGT 1344
QY 2392 TACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACACACCGCTGGTAGCGG 2451
Db 1343 TACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACACACCGCTGGTAGCGG 1284
QY 2452 TGGTTTTTTTGTGCAAGCAGCAGATTAACGGCAGAAAAAAGATCTCAAGAAGATCC 2511
Db 1283 TGGTTTTTTTGTGCAAGCAGCAGATTAACGGCAGAAAAAAGATCTCAAGAAGATCC 1224
QY 2512 TTTGATCTTTTCTACGGGCTCTGACCTCAGTGGACGAAACTCAGTTAAGGATTTT 2571
Db 1223 TTTGATCTTTTCTACGGGCTCTGACCTCAGTGGACGAAACTCAGTTAAGGATTTT 1164
QY 2572 GGTCTAGATTAATCAAAAAGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTT 2631
Db 1163 GGTCTAGATTAATCAAAAAGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTT 1104
QY 2632 TAAATCAATCTAAAGTATATATAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAG 2691
Db 1103 TAAATCAATCTAAAGTATATATAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAG 1044
QY 2692 TGAGGCACCTATCTACGCATCTGCTATTTTGGTTCATCATAGTTGCGCTGACTCCCGT 2751
Db 1043 TGAGGCACCTATCTACGCATCTGCTATTTTGGTTCATCATAGTTGCGCTGACTCCCGT 984
QY 2752 CGTGTAGATAAATACATACGAGGAGGCTTACCATCTGCGCCCGAGTGTGCAATGATACC 2811
Db 983 CGTGTAGATAAATACATACGAGGAGGCTTACCATCTGCGCCCGAGTGTGCAATGATACC 924
QY 2812 GCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCGGAGGCG 2871
Db 923 GCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCGGAGGCG 864
QY 2872 CGAGCGCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGGTCCG 2931
Db 863 CGAGCGCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGGTCCG 804
QY 2932 GGAAGCTAGAGTAAGTAGTTCGCGAGTAAATAGTTTGGCAAGTGTGGCCATGCTACTAC 2991
Db 803 GGAAGCTAGAGTAAGTAGTTCGCGAGTAAATAGTTTGGCAAGTGTGGCCATGCTACTAC 744
QY 2992 AGGCATCGTGGTGTACAGCTCGTGGTATGGTTCATTACAGCTCCGGTTCCCAAGC 3051
Db 743 AGGCATCGTGGTGTACAGCTCGTGGTATGGTTCATTACAGCTCCGGTTCCCAAGC 684
QY 3052 ATCAAGCGCAGTACATGATCCCGCATGTTGTGCAAAAAGCGGTTAGCTCTCTGGTCC 3111
Db 683 ATCAAGCGCAGTACATGATCCCGCATGTTGTGCAAAAAGCGGTTAGCTCTCTGGTCC 624
QY 3112 TCCGATCGTGTGTCAGAAAGTAAGTTGGCCGAGTTTATCACTCATGTTATGCGACACT 3171
Db 623 TCCGATCGTGTGTCAGAAAGTAAGTTGGCCGAGTTTATCACTCATGTTATGCGACACT 564
QY 3172 GCATAATCTCTTACTGTATGTCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTACTC 3231
Db 563 GCATAATCTCTTACTGTATGTCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTACTC 504
QY 3232 AACCAAGTCAATCTGAGAAATAGTGTATGCGGACCGAGTTGCTCTTGGCCCGCGTCAAT 3291
Db 503 AACCAAGTCAATCTGAGAAATAGTGTATGCGGACCGAGTTGCTCTTGGCCCGCGTCAAT 444
QY 3292 ACGGGATAATACCGGCCACATAGCAGAACTTTTAAAGTGCTCATCTTTGGAAAAAGCTTC 3351
Db 443 ACGGGATAATACCGGCCACATAGCAGAACTTTTAAAGTGCTCATCTTTGGAAAAAGCTTC 384
```

```
QY 3352 TTCCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCAC 3411
Db 383 TTCCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCAC 324
QY 3412 TCGTGCAACCAACTGATCTTTCAGCATCTTTTACTTTTCCAGCGTTTCTGGGTGAGCAAA 3471
Db 323 TCGTGCAACCAACTGATCTTTCAGCATCTTTTACTTTTCCAGCGTTTCTGGGTGAGCAAA 264
QY 3472 AACAGGAAGCGCAAAATGCGCAAAAAGGAATAAGGGCGACACGGAAATGTTGAATACT 3531
Db 263 AACAGGAAGCGCAAAATGCGCAAAAAGGAATAAGGGCGACACGGAAATGTTGAATACT 204
QY 3532 CATACTCTTCTTCTTCAATATATTGAAGCATTTATCAGGGTTATGTTCTCATGAGCGG 3591
Db 203 CATACTCTTCTTCTTCAATATATTGAAGCATTTATCAGGGTTATGTTCTCATGAGCGG 144
QY 3592 ATACATATTGTAATGTTATTTAGAAAAATAAACAATAGGGTTCCCGCGACATTTCCCGG 3651
Db 143 ATACATATTGTAATGTTATTTAGAAAAATAAACAATAGGGTTCCCGCGACATTTCCCGG 84
QY 3652 AAAAGTGCCAC 3662
Db 83 AAAAGTGCCAC 73

RESULT 14
CQ802124/c
LOCUS CQ802124 3918 bp DNA linear PAT 05-MAY-2004
DEFINITION Sequence 109 from Patent WO2004007717.
ACCESSION CQ802124
VERSION CQ802124.1 GI:47058654
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Unger, C.M., Beste, G., Jensen, K.H., Zehetmeier, C., Knauer, R.,
Guelbenzu, B.L., Torella, C. and Ilag, L.L.
TITLE Proteomic screen to identify disease-related biological
molecules and inhibitors thereto
JOURNAL Patent: WO 2004007717-A 109 22-JAN-2004;
Xerion Pharmaceuticals AG (DE)
FEATURES
source
1..3918
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description of Artificial Sequence: vector pXP10"
ORIGIN

Query Match 68.8%; Score 2517.8; DB 6; Length 3918;
Best Local Similarity 81.8%; Pred. No. 0;
Matches 3052; Conservative 0; Mismatches 537; Indels 142; Gaps 7;

QY 3 AAATTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGTCTCAT 62
Db 3732 AAATTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGTCTCAT 3673
QY 63 TTTTAAACCAATAGCGCGAAATCGCAAAATCCCTTATAAATCAAAAGAAATACACCGAGA 122
Db 3672 TTTTAAACCAATAGCGCGAAATCGCAAAATCCCTTATAAATCAAAAGAAATACACCGAGA 3613
QY 123 TAGGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGACTCCA 182
Db 3612 TAGGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGACTCCA 3553
QY 183 AGCTAAAGGGCGAAAAACCGTCTATCAGGGCGATGCGCCACTACGTGAACCATCACCT 242
Db 3552 AGCTAAAGGGCGAAAAACCGTCTATCAGGGCGATGCGCCACTACGTGAACCATCACCA 3493
QY 243 AATCAAGTTTTTTGGGTCGAGTCCGTAAGACACTTAATCGGAACCTTAAAGGAGGCC 302
Db 3492 AATCAAGTTTTTTGGGTCGAGTCCGTAAGACACTTAATCGGAACCTTAAAGGAGGCC 3433
```

QY	303	CCCAGTTTAGAGCTTTGACGGGAAAGCCGGCGAAAGCTGCGGAGAAAGGAAGGAAGAAG	362	QY	1312	TCCGATGGTCAGCCCGAAGCGGCATCAGCAACCCGAAACAATACCGGCGACACCGCGAACT	1371
Db	3432	CCCAGTTTAGAGCTTTGACGGGAAAGCCGGCGAAAGCTGCGGAGAAAGGAAGGAAGAAG	3373	Db	2379	AGAGCCACTCCGCTCGAACCAGCCCTCCACCTGTGCA	2344
QY	363	CGAAAGGAGCGGCGCTAGGCGCTGGCAAGTGTAGCGGTCAACGCTGCGGTAAACCA	422	QY	1372	GCCGTGCCGTGTGCAGATTAAATGACAGGGTGCAGCGCTGGGATATTACGTACGGAGG	1431
Db	3372	CGAAAGGAGCGGCGCTAGGCGCTGGCAAGTGTAGCGGTCAACGCTGCGGTAAACCA	3313	Db	2343	-----CTGGACCTGGGCCATGGCGCTGGCGCGCATAGAAAGGAACAACATAAGGAA	2291
QY	423	CACCGCGCGCTTAATGGCGCGCTACAGGGCGCT	458	QY	1432	ACGGGTATCCTGGCTGGATGCGCAGAAATGACATGGATACCCCGTGAGTTACCGCGG	1491
Db	3312	CACCGCGCGCTTAATGGCGCGCTACAGGGCGCTACTATGGTTGCTTTGACGTATCGG	3253	Db	2290	TTGCCAATAAATTTTTCAGTTGAAATCTCAAAAAA	2244
QY	459	-----CCATTTCGCCATT	471	QY	1492	GCGCGCTTGGCGTAATCATGTTCATAGCTGTTCCTGTGTGAAAATTGTATCCGCTCAC	1551
Db	3252	GTGTGAATACCGCACAGATGCGTAAGGAGAAATACCGCATCAGCGCGCATTCGCCATT	3193	Db	2243	CCAAAGCTTGGCGTAATCATGTTCATAGCTGTTCCTGTGTGAAATTTGTATCCGCTCAC	2184
QY	472	CAGGTGCGCAACTGTTGGGAAGGCGATCGGTGCGGCGCTTTCGCTATTACGCCAGCT	531	QY	1552	AATTCCACAACAATACAGCGCGGAAGCATAAAGTGTAAAGCTGGGTGCTTAATGAGT	1611
Db	3192	CAGGTGCGCAACTGTTGGGAAGGCGATCGGTGCGGCGCTTTCGCTATTACGCCAGCT	3133	Db	2183	AATTCCACAACAATACAGCGCGGAAGCATAAAGTGTAAAGCTGGGTGCTTAATGAGT	2124
QY	532	GGGAAAGGGGATGTGCTGCAAGCGGATTAAGTTGGGTAAAGCGCAGGGTTTCCAGTC	591	QY	1612	GAGCTAACTCACTTAATTTGCGCTCACTGCGCGCTTCCAGTGGGAAACCTGTC	1671
Db	3132	GGGAAAGGGGATGTGCTGCAAGCGGATTAAGTTGGGTAAAGCGCAGGGTTTCCAGTC	3073	Db	2123	GAGCTAACTCACTTAATTTGCGCTCACTGCGCGCTTCCAGTGGGAAACCTGTC	2064
QY	592	ACGAGCTTGTAAACGACGCGCAGTGAAGCGCGCTCGTTTCACTGCTTTTGAACCCGT	651	QY	1672	GTGCCAGCTGCAATTAATGAATCGGCCAACGCGCGGGAGAGCGGTTTTCGCTATTGGCG	1731
Db	3072	ACGAGCTTGTAAACGACGCGCAGTGAAGCGCGCTCGTTTCACTGCTTTTGAACCCGT	3013	Db	2063	GTGCCAGCTGCAATTAATGAATCGGCCAACGCGCGGGAGAGCGGTTTTCGCTATTGGCG	2004
QY	652	GGAGGACGGGAGACTCGCGGTGCAAAATGTGTTTACAGCGGTGATGGAGCAGATGAAGT	711	QY	1732	CTCTTCGCGCTTCTCGCTCACTGACTCGCTGGGCTCGGTTCGCTGCGGCGAGCGGT	1791
Db	3012	-----TAGCAACGTCGAAATATACATATAAAGGTGGCAACATATAAAGAA	2964	Db	2003	CTCTTCGCGCTTCTCGCTCACTGACTCGCTGGCTCGGTTCGCTGCGGCGAGCGGT	1944
QY	712	GCTCGACAGCTGCAGACACGACACTAGATTAAACCTAGAAAGTATCATATTGTGAC	771	QY	1792	ATCAGCTCACTCAAAAGCGGTAAATACGGTTTATCCACAGAAATCAGGGGATTAACGCAAGGAA	1851
Db	2963	CGCAAGACACACAGGAAATAGTTTATTTGTGCACAATCAATAGAAATTCATATGGTTT	2904	Db	1943	ATCAGCTCACTCAAAAGCGGTAAATACGGTTTATCCACAGAAATCAGGGGATTAACGCAAGGAA	1884
QY	772	GTACGTTAAAGATATCATGCGTAAATTTGACGCATGGATCTGTAATACGACTCACTAT	831	QY	1852	GAACATGTGAGCAAAAGCCAGCAAAAGCCAGGAAACCGTAAAAAGGCGCGGTTCTCTGCG	1911
Db	2903	ACCAGCGCAAGACATTAAGGCGACATTCACCGACTGAGGCAAGAGGTAAATATTG	2844	Db	1883	GAACATGTGAGCAAAAGCCAGCAAAAGCCAGGAAACCGTAAAAAGGCGCGGTTCTCTGCG	1824
QY	832	AGGGCAATGGGTACCGGGCCCCCTCGAGGTGCGAGGTGCGATCGATTAAGCTTGTATCG	891	QY	1912	GTTCCTCATAGGCTCCGCGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAG	1971
Db	2843	ACGGAAATTTATTAATAAGGTGAATTTATCACCGTCAACCGACTTGAGCAATTTGGGAAT	2784	Db	1823	GTTCCTCATAGGCTCCGCGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAG	1764
QY	892	AATTCCTCAGCGCGGGATCCACTAGTTCTAGAGCGCGCCACCGCGGTGGAGCTCC	951	QY	1972	GTGGGAAACCCGACAGGACTATAAGATACAGCGGCTTCCCGCTGGAAGCTCCCTCGT	2031
Db	2783	AGAGCCAGCA-----AATCACAGTAGCACCAATTACCAATTAGCAAGGCGCGAAACG	2732	Db	1763	GTGGGAAACCCGACAGGACTATAAGATACAGCGGCTTCCCGCTGGAAGCTCCCTCGT	1704
QY	952	AGCTTTTGTTCCTTTAGTGAGGTTAATTAGATCCCATGGTCAATTTTACGCAAGACTA	1011	QY	2032	GCGCTCTCCTGTCCGACCCCTGCGGATACCGGATACCGCTGCGGCTTCTCCCTTCGGG	2091
Db	2731	TCACCAATGAACCATCATAGCAGCACCGTAAATCAGTAGCAGACAGATCA---AGTTTG	2675	Db	1703	GCGCTCTCCTGTTCGACCCCTGCGGATACCGGATACCGCTGCGGCTTCTCCCTTCGGG	1644
QY	1012	TCTTTCTAGGTTAATCTAGCTGCAATCAGGATCATATCGTGGGTCTTTTTTTCGGCTCA	1071	QY	2092	AAGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGCGGTGAGTGTGCTGTCG	2151
Db	2674	CCTTTAGCGTCAGACTGAGCGGCTTTTCATCGGCAITTTTCGCTCATAGCCCTTTATTA	2615	Db	1643	AAGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGCGGTGAGTGTGCTGTCG	1584
QY	1072	GTCAATCGCCAGCTGGCGCTATCTGGCATCGGGAGGAGAGCCCGTGCCTTTTCC	1131	QY	2152	CTCAAGCTGGCGCTGTGTGACCAACCCCGCTTTCAGCGCCAGCCGCTGCGCTTATCCGG	2211
Db	2614	GCGTTTGCATTTTTCATATAATCAAAATCACCGAAACCGAGCGCGCCACCGGAACCGCCA	2555	Db	1583	CTCAAGCTGGCGCTGTGTGACCAACCCCGCTTTCAGCGCCAGCCGCTGCGCTTATCCGG	1524
QY	1132	GCGAGGTTGAAGCGGCATGGAAGAGTTTGGCGAGGATGACTGTGCTGCAATGACGTTG	1191	QY	2212	TAACTATCGTTGAGTCCAAACCCGGTAAAGACAGCTTATCGCCACTGGCAGCAGCCAC	2271
Db	2554	CCCTCAGAGCGCCACCTCAGAACCGCACCTCAGAGCGCCACCTCAGAGCGCCA	2495	Db	1523	TAACTATCGTTGAGTCCAAACCCGGTAAAGACAGCTTATCGCCACTGGCAGCAGCCAC	1464
QY	1192	AGCGAAACCGCATGATTAATTCAGTATGATTTGGGAAGGTGTGGCCATGACGCGCTTAAAG	1251	QY	2272	TGTTAAGAGGATTAAGCAGAGCGGTATGATAGGCGGTCTACAGAGTCTTGAAGTGTG	2331
Db	2494	CGAGAACCAACAC-----CAGAGCGCGCGCCAGCTATCGGCGACGCGGTTCCAGCGAT	2440	Db	1463	TGTTAAGAGGATTAAGCAGAGCGGTATGATAGGCGGTCTACAGAGTCTTGAAGTGTG	1404
QY	1252	GTGAACTGTTTCAGGCGCACTGGGATACAGTTTCGTCGCGGCTTTTCGCGACACAGT	1311	QY	2332	GCCTAACTACCGCTACACTAGAAAGCAGATTTTGGTATCTGCGCTCTGCTGAAGCCAGT	2391
Db	2439	CCGGATACGGCACCGCGCGCACTTCGCGCGCGGCTTTGATGAGTGCATCTTCGCGCACCGCC	2380	Db	1403	GCCTAACTACCGCTACACTAGAAAGCAGATTTTGGTATCTGCGCTCTGCTGAAGCCAGT	1344
				QY	2392	TACCTTCGGAAGAAAGAGTTGGTAGCTTTTGTATCCGGGCAACAAACCAACCGCTGGTAGCGG	2451

1343	Db	TACCTTCGGAAGAAGCTGGTAGCTCTTGATCCGGCAAAACAAACCAACCGCTGGTAGCGG	1284
2452	Qy	TGGTTTTTTTTTGTTCGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCC	2511
1283	Db	TGGTTTTTTTTTGTTCGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCC	1324
2512	Qy	TTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACCTACGTTAAAGGATTTT	2571
1223	Db	TTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACCTACGTTAAAGGATTTT	1164
2572	Qy	GGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCGTTTAAATTAATAATGAAGTTT	2631
1163	Db	GGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCGTTTAAATTAATAATGAAGTTT	1104
2632	Qy	TAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTTACCAATGCTTAATCAG	2691
1103	Db	TAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTTACCAATGCTTAATCAG	1044
2692	Qy	TGAGGCACCTATCTCAGCGATCTGTCTATTCTGCTTCATCATAGTTGCCTGACCTCCCGT	2751
1043	Db	TGAGGCACCTATCTCAGCGATCTGTCTATTCTGCTTCATCATAGTTGCCTGACCTCCCGT	984
2752	Qy	CGTGTAGATAACTACGATACGGAGGGCTTACCATCTGGCCCCAGTGCTCGAATGATACC	2811
983	Db	CGTGTAGATAACTACGATACGGAGGGCTTACCATCTGGCCCCAGTGCTCGAATGATACC	924
2812	Qy	GCAGAGCCACGCTCACCGGCTCCAGATTTATCAGCAATAAAACAGCCAGCCCGAAGGGC	2871
923	Db	GCAGAGCCACGCTCACCGGCTCCAGATTTATCAGCAATAAAACAGCCAGCCCGAAGGGC	864
2872	Qy	CGAGCGCAGAAGTGTCTCTCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGTGCCG	2931
863	Db	CGAGCGCAGAAGTGTCTCTCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGTGCCG	804
2932	Qy	GGAACTTAGAGTAAAGTAGTTTCGCCAGTTAATAGTTTCGGCAAAGTTGTTGCCAATTGCTAC	2991
803	Db	GGAACTTAGAGTAAAGTAGTTTCGCCAGTTAATAGTTTCGGCAAAGTTGTTGCCAATTGCTAC	744
2992	Qy	AGGCATCGTGGTGTACGCTCGTCTGTTGTGATGCGTTCATCAGCTCCGGTCCCAACG	3051
743	Db	AGGCATCGTGGTGTACGCTCGTCTGTTGTGATGCGTTCATCAGCTCCGGTCCCAACG	684
3052	Qy	ATCAAGCGCAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTAGCTCCTTCGGTCC	3111
683	Db	ATCAAGCGCAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTAGCTCCTTCGGTCC	624
3112	Qy	TCCGATCGTTGTGAGAAAGTTGGCGCAGTGTGTTATCATCTCATGGTTATGGCAGCACT	3171
623	Db	TCCGATCGTTGTGAGAAAGTTGGCGCAGTGTGTTATCATCTCATGGTTATGGCAGCACT	564
3172	Qy	GCATAATTCCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGTCAGTACTC	3231
563	Db	GCATAATTCCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGTCAGTACTC	504
3232	Qy	AACCAAGTCATTCTGAGAAAGTGTATGCGCGCACCGAGTTGCTTCTTCCCGCGCTCAAT	3291
503	Db	AACCAAGTCATTCTGAGAAAGTGTATGCGCGCACCGAGTTGCTTCTTCCCGCGCTCAAT	444
3292	Qy	ACGGGATAATACCGCGCCAATAGCAGAACTTTAAAAAGTGCTCATCATGAAACGTTTC	3351
443	Db	ACGGGATAATACCGCGCCAATAGCAGAACTTTAAAAAGTGCTCATCATGAAACGTTTC	384
3352	Qy	TTCCGGCGGAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCCAC	3411
383	Db	TTCCGGCGGAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCCAC	324
3412	Qy	TCGTGCACCAACTGATCTTTCAGCATCTTTTACTTTTCCAGCGGTTTCTGGGTGAGCAAA	3471
323	Db	TCGAGCACCCAACTGATCTTTCAGCATCTTTTACTTTTCCAGCGGTTTCTGGGTGAGCAAA	264
3472	Qy	AACAGGAAGCAAAATGCCGCAAAAAAGGAAATAAGGGCGACACGGAATGTTGAATACT	3531

Db	263	AACAGGAAGGCAAAATGCCGCCCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACT	204
Qy	3532	CATACTCTTCTCTTTTTCAAATATTATTGAAGCACTTATCAGGGTATTGTCTCATGAGCGG	3591
Db	203	CATACTCTTCTCTTTTTCAAATATTATTGAAGCACTTATCAGGGTATTGTCTCATGAGCGG	144
Qy	3592	ATACATATTGTAATGTATTTAGAAAAATAAACAATAGGGGTTCCGGCGACATTTCCCGG	3651
Db	143	ATACATATTGTAATGTATTTAGAAAAATAAACAATAGGGGTTCCGGCGACATTTCCCGG	84
Qy	3652	AAAAGTCCAC	3662
Db	83	AAAAGTCCAC	73
RESULT 15			
CQ873225/c			
LOCUS	CQ873225	3918 bp	DNA linear PAT 27-SEP-2004
DEFINITION	Sequence 3 from Patent EP1457499.		
ACCESSION	CQ873225		
VERSION	CQ873225.1 GI:52746953		
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 Jay, D.G., Eustace, B.K., Sakurai, T. and Beste, G.		
TITLE	Inhibitors of extracellular Hsp90		
JOURNAL	Patent: EP 1457499-A 3 15-SEP-2004;		
	TUFTS UNIVERSITY SCHOOL OF MEDICINE (US); Xerion Pharmaceuticals AG (DE)		
FEATURES	Location/Qualifiers		
source	1..3918		
	/organism="Mus musculus"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:10090"		
ORIGIN			
Query Match	68.8%	Score 2517,8;	DB 6; Length 3918;
Best Local Similarity	81.8%	Pred. No. 0;	
Matches 3052;	Conservative	0; Mismatches 537;	Indels 142; Gaps 7;
Qy	3	AAATTGTAAGCGTTAATATTTCGTAAAAATTCGCGTTAAATTTTGTAAATCAGTCTCAT	62
Db	3732	AAATTGTAAGCGTTAATATTTCGTAAAAATTCGCGTTAAATTTTGTAAATCAGTCTCAT	3673
Qy	63	TTTTTAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGCCGAGA	122
Db	3672	TTTTTAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGCCGAGA	3613
Qy	123	TAGGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGACGTGGACTCCA	182
Db	3612	TAGGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGACGTGGACTCGA	3553
Qy	183	ACGTCAAAAGGCGGAAACCCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCT	242
Db	3552	ACGTCAAAAGGCGGAAACCCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCA	3493
Qy	243	AATCAAGTTTTTGGGGTTCGAGTGGCGTTAAGCACTAAATCGGAACCCCTAAAGGGAGCC	302
Db	3492	AATCAAGTTTTTGGGGTTCGAGTGGCGTTAAGCACTAAATCGGAACCCCTAAAGGGAGCC	3433
Qy	303	CCGATTTTACAGCTTCGACGGGGAAGCCGCGAAGCTGGCGAGAAAGGAGGGAAGAAAG	362
Db	3432	CCGATTTTACAGCTTCGACGGGGAAGCCGCGAAGCTGGCGAGAAAGGAGGGAAGAAAG	3373
Qy	363	CGAAAGGAGCGGCGCTAGGCGCGTGCAGAGTGTAGCGGTACGCTCGCGTAAACCA	422
Db	3372	CGAAAGGAGCGGCGCTAGGCGCGTGCAGAGTGTAGCGGTACGCTCGCGTAAACCA	3313
Qy	423	CACCCGCGCGCTTAATGCGCGCTACAGGGCGGT	458

Db 3312 CACCCGCGGGCTTAATGCGCGCTACAGGGCGCTACTATGTTGCTTTGACGTATGCG 3253
Qy 459 -----CCCATTCGCCATT 471
Db 3252 GTGTGAATATACCGCACAGATGCGTAAGGAGAAAAATACCGCATCAGGGCCATTCGCCATT 3193
Qy 472 CAGGCTGGCCAACTGTGTGGAAAGGCGATCGGTGGGGCCTCTTCGCTATTACGCCAGCT 531
Db 3192 CAGGCTGGCCAACTGTGTGGAAAGGCGATCGGTGGGGCCTCTTCGCTATTACGCCAGCT 3133
Qy 532 GCGAAAGGGGATGTGCTGCAAGCGGATTAAGTTGGGTAAGCCAGGTTTTCAGTC 591
Db 3132 GCGAAAGGGGATGTGCTGCAAGCGGATTAAGTTGGGTAAGCCAGGTTTTCAGTC 3073
Qy 592 ACGAGCTTTGTAAGACGAGCGGCAGTGAGCGGCTCGTTTCATTCAGCTTTTGAACCCGT 651
Db 3072 ACGAGCTTTGTAAGACGAGCGGCAGTGGAATCTTATTAAGACTCTTATACGCCATATGT 3013
Qy 652 GGAGGACGGGAGACTCGCGGTGCAAAATGTGTTTACAGCGTGATGGAGCAGATGAAGAT 711
Db 3012 -----TAGCAAGCTCGAAATATACATATAAGGTGGCAACATATAAGAAA 2964
Qy 712 GCTGACAGCTGCGAGAACACGCGACTAGATTAACTCTAGAAAGATTAATCATATTGTGAC 771
Db 2963 CGAAAGACACACCGGAATAAGTTTATTTGTCACAATCAATAGAAAAATTCATATGGTTT 2904
Qy 772 GTACGTTAAGATATCATGCTAAATTTGACGCATGGGATCTGTAATACGACTCACTAT 831
Db 2903 ACCAGCGCAAGACATAGGGCGCATTTCAACGACTGAGGCAAGAGGTAAATATTG 2844
Qy 832 AGGGCGAAATGGGTACCGGGCCCCCTCGAGTGCAGCGGTATCGATAAGCTTTGATATCG 891
Db 2843 ACGGAATATTTCATTAAGGTGAATATACCGCTCAGCGCTTGAGCCATTGGGAATT 2784
Qy 892 AATTCCTGAGCGCGGGGATCCATAGTTCTAGAGCGGCGGCCACCGGGTGGAGCTCC 951
Db 2783 AGAGCCAGCA-----AAATCACCAGTAGCACCAATTACCAATTAGCAAGGCGGAAAACG 2732
Qy 952 AGCTTTTGTTCCTTTAGTGAGGGTTAATTAGATCCCATCGCTCAATTTTACGACAGCTA 1011
Db 2731 TCACCAATGAACCACTGATAGCAGCACCGTAACTAGTAGCAGACAGATCA---AGTTTG 2675
Qy 1012 TCTTTCTAGGGTTAATCTAGCTGCATCAGGATCATATCGTGGGTCTTTTTCGGCTCA 1071
Db 2674 CTTTATAGCTCAGACTGTAGCGGTTTTCATCGGCAATTTTCGTTATAGCCCTTATTA 2615
Qy 1072 GTCATCGCCAGCTGGCGCTATCTGGGCATCGGGAGGAGAACCGCGTCTTTTCCC 1131
Db 2614 GCGTTTGCCATTTTTCATATAATCAAAATCACCGGAAACCGGAGCCGCCACCGGAAACGCGCA 2555
Qy 1132 GCGAGTTTGAAGCGCATGGAAGAGTTTGGCGAGGATGACTGCTGCTGCAATTGACGTTG 1191
Db 2554 CCTCAGAGCCGCCACCTCAGAACCGGCACCTCAGAGCCGCCACCTCAGAGCGGCA 2495
Qy 1192 AGCGAAAAAGCGCTTTTACCATGATGATTGGGAAAGGTGTGGCCATGCAAGCTTTTAAAG 1251
Db 2494 CCAGAAACCAAC-----CAGAGCGCGCGCAGCTATGCGGCAAGCGGTTCCAGCGGAT 2440
Qy 1252 GTGAATGTTCTGTTAGGCCCACTGGGATACAGTTTCTGCGGCTTTTTCGGGACACAGT 1311
Db 2439 CCGGATAGCGCACCGCGCGCACCTGGGGCGCGGTTTGTATGATGCACTTTCCGCCACCGCC 2380
Qy 1312 TCCGGATGTCAGCGGAGCGCATCAGCAACCGGAAACATACCGGCGACAGCGGAACT 1371
Db 2379 AGAGCCACCTTCGCTGTAACCGGCTCCACTGTGCA----- 2344
Qy 1372 GCGCTGCCGCTGTGAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTACGCGAGG 1431
Db 2343 -.-.-.-.-CTGGACCTGGGCCATGGCGGCTGGCGCGCATAGAAAGAAACAACTAAAGGAA 2291
Qy 1432 ACGGATATCTCGCTGGATGCGCGAGAAATGACATGGAATACCCGTTAGTTACCGGG 1491
Db 2290 TTGCGAATAATAATTTTTCACGTTTGAATAATCTCCAAAAA-----GGCT 2244

Qy 1492 GCGCGCTTGGCGTAATCATGCTCATAGCTGTTTCTCTGTGAAATTTGTTATCCGCTCAC 1551
Db 2243 CCAAGCTTGGCGTAATCATGCTCATAGCTGTTTCTCTGTGAAATTTGTTATCCGCTCAC 2184
Qy 1552 AATTCACACAAACATACAGAGCGGGAAGCATAAAGTGTAAAGCTGGGTGCTTAATGAT 1611
Db 2183 AATTCACACAAACATACAGAGCGGGAAGCATAAAGTGTAAAGCTGGGTGCTTAATGAT 2124
Qy 1612 GAGCTAACTCAATTAATTTGCTGCGCTCACTGCCCGCTTCCAGTGGGAAACCTGTC 1671
Db 2123 GAGCTAACTCAATTAATTTGCTGCGCTCACTGCCCGCTTCCAGTGGGAAACCTGTC 2064
Qy 1672 GTGCCAGCTGCAATTAATGAATCGGCAACCGCGCGGAGAGCGGTGTTGCGTATTGGGCG 1731
Db 2063 GTGCCAGCTGCAATTAATGAATCGGCAACCGCGCGGAGAGCGGTGTTGCGTATTGGGCG 2004
Qy 1732 CTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTGCTGCTGCGGCGAGCGGT 1791
Db 2003 CTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTGCTGCTGCGGCGAGCGGT 1944
Qy 1792 ATCAGCTCACTCAAGGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGAGGAA 1851
Db 1943 ATCAGCTCACTCAAGGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGAGGAA 1884
Qy 1852 GAAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAACCGTAAAAAGGCGCGCTTGTGCG 1911
Db 1883 GAAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAACCGTAAAAAGGCGCGCTTGTGCG 1824
Qy 1912 GTTTTTCATAGGCTCCGCCCCCTGACAGCATCACAAAAATCGACGCTCAAGTCAGAG 1971
Db 1823 GTTTTTCATAGGCTCCGCCCCCTGACAGCATCACAAAAATCGACGCTCAAGTCAGAG 1764
Qy 1972 GTGGCGAAACCGCAGAGCATATAAGATACAGAGCTTTTCCCTCGGAAGCTCCCTCGT 2031
Db 1763 GTGGCGAAACCGCAGAGCATATAAGATACAGAGCTTTTCCCTCGGAAGCTCCCTCGT 1704
Qy 2032 GCGCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTGCTCGCTTTTCTCCCTTCGGG 2091
Db 1703 GCGCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTGCTCGCTTTCTCCCTTCGGG 1644
Qy 2092 AAGCTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGCTGCTGAGTGGTTG 2151
Db 1643 AAGCTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGCTGCTGAGTGGTTG 1584
Qy 2152 CTCCAAGCTGGGCTGTGTGCAGAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGG 2211
Db 1583 CTCCAAGCTGGGCTGTGTGCAGAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGG 1524
Qy 2212 TAACTATCTGTTGAGTCCAACCCCGTAAAGACAGACTTATCGCCACTGGCAGAGCCAC 2271
Db 1523 TAACTATCTGTTGAGTCCAACCCCGTAAAGACAGACTTATCGCCACTGGCAGAGCCAC 1464
Qy 2272 TGGTAAAGATTAGCAGAGGAGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGT 2331
Db 1463 TGGTAAAGATTAGCAGAGGAGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGT 1404
Qy 2332 GCCTAACTACGGTACACTAGAGGACAGTATTTGCTATCTGCGCTCTGCTCAAGCCAGT 2391
Db 1403 GCCTAACTACGGTACACTAGAGGACAGTATTTGCTATCTGCGCTCTGCTCAAGCCAGT 1344
Qy 2392 TACCTTCGGAAAAAGATTGGTAGCTCTTGTATCCGGCAAAACAAACCGCTGGTAGCGG 2451
Db 1343 TACCTTCGGAAAAAGATTGGTAGCTCTTGTATCCGGCAAAACAAACCGCTGGTAGCGG 1284
Qy 2452 TGGTTTTTTTGTTCGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCC 2511
Db 1283 TGGTTTTTTTGTTCGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCC 1224
Qy 2512 TTTGATCTTTTCTACGGGCTGACGCTCAGTGGAAACGAAAACTCACGTTAAGGGATTTT 2571
Db 1223 TTTGATCTTTTCTACGGGCTGACGCTCAGTGGAAACGAAAACTCACGTTAAGGGATTTT 1164

QY	2572	GGTCATGAGATTATCAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAATGAAGTTT	2631
Db	1163	GGTCATGAGATTATCAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAATGAAGTTT	1104
QY	2632	TAAATCAATCTAAAGTATATATAGTAAACTTGGTCTGACAGATTACCAATGCTTAATCAG	2691
Db	1103	TAAATCAATCTAAAGTATATATAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAG	1044
QY	2692	TGAGGCACTTATCTCAGCGATCTGTCTATTTCTGTTTATCCATAGTTGCCCTGACTCCCCGT	2751
Db	1043	TGAGGCACTTATCTCAGCGATCTGTCTATTTCTGTTTATCCATAGTTGCCCTGACTCCCCGT	984
QY	2752	CGTGTAGATAACTACGATCGGAGGGCTTACCATCTGGCCCCAGTCTGCAATGATACC	2811
Db	983	CGTGTAGATAACTACGATCGGAGGGCTTACCATCTGGCCCCAGTCTGCAATGATACC	924
QY	2812	GCAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGC	2871
Db	923	GCAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGC	864
QY	2872	CGAGCCGAGAAGTGGTCTTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGGCCG	2931
Db	863	CGAGCCGAGAAGTGGTCTTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGGCCG	804
QY	2932	GGNAGCTAGAGTAAGTATGTTCCGCACTTAATAGTTTCCGCAACGTTTGGCCATTCCTAC	2991
Db	803	GGNAGCTAGAGTAAGTATGTTCCGCACTTAATAGTTTCCGCAACGTTTGGCCATTCCTAC	744
QY	2992	AGGCATCTGGTGTACAGCTCGTCTGTTGGTATGGCTTCAATTCAGCTCCGGTTCCCAACG	3051
Db	743	AGGCATCTGGTGTACAGCTCGTCTGTTGGTATGGCTTCAATTCAGCTCCGGTTCCCAACG	684
QY	3052	ATCAAGCCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCC	3111
Db	683	ATCAAGCCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCC	624
QY	3112	TCCGATCTGTTGTCAGAAAGTATGGCCGAGTGTATCACTCATGTTATGSCAGCACT	3171
Db	623	TCCGATCTGTTGTCAGAAAGTATGGCCGAGTGTATCACTCATGTTATGSCAGCACT	564
QY	3172	GCATAATCTCTTACTGTATGTCATCCGATCCGATGATGTTTCTGTCAGTGGTGAGTACTC	3231
Db	563	GCATAATCTCTTACTGTATGTCATCCGATCCGATGATGTTTCTGTCAGTGGTGAGTACTC	504
QY	3232	AACCAAGTCAATCTGAGAATAGTGTATGCGGCGACCGAGTGTCTTTGCCCGCGTCAAT	3291
Db	503	AACCAAGTCAATCTGAGAATAGTGTATGCGGCGACCGAGTGTCTTTGCCCGCGTCAAT	444
QY	3292	ACGGGATATACCGCGCCACATAGCAGAACTTTAAAGTCTCATCATTTGGAAAAAGTTTC	3351
Db	443	ACGGGATATACCGCGCCACATAGCAGAACTTTAAAGTCTCATCATTTGGAAAAAGTTTC	384
QY	3352	TTCCGGGGGGAAGAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCAC	3411
Db	383	TTCCGGGGGGAAGAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCAC	324
QY	3412	TCGTGACACCCAACTGATCTTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAA	3471
Db	323	TCGTGACACCCAACTGATCTTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAA	264
QY	3472	AACAGGAAGGCAAAATGCGCGCAAAAAAGGAATAAGGGCGACACGGAAATGTTGAATACT	3531
Db	263	AACAGGAAGGCAAAATGCGCGCAAAAAAGGAATAAGGGCGACACGGAAATGTTGAATACT	204
QY	3532	CATACCTCTCTTTTCAATATTTTGAAGCATTTTATCAGGGTTATGCTCATGAGCGG	3591
Db	203	CATACCTCTCTTTTCAATATTTTGAAGCATTTTATCAGGGTTATGCTCATGAGCGG	144
QY	3592	ATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCGG	3651
Db	143	ATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCGG	84
QY	3652	AAAAAGTGCCAC	3662

Db 83 AAAAGTGCCAC 73

Search completed: June 5, 2005, 19:22:50
Job time : 15400.2 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 13:34:58 ; Search time 1772.75 Seconds
(without alignments)
12228.534 Million cell updates/sec

Title: US-10-826-523-41

Perfect score: 3662

Sequence: 1 ccaatttaagcgtaata.....atttccccgaaagtgcac 3662

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3648.4	99.6	3661	10	ACA55354
2	2665.4	72.8	3637	11	ADM68463
3	2665.4	72.8	3637	12	ADP26616
4	2665.4	72.8	3637	13	ADQ88438
5	2665.4	72.8	3637	13	ADQ88438
6	2664.8	72.8	3637	8	ABX14571
7	2664.8	72.8	3637	11	ADM68448
8	2664.8	72.8	3637	12	ADP26601
9	2664.8	72.8	3637	13	ADQ88423
10	2664.8	72.8	3637	13	ADQ88423
11	2656.4	72.5	3637	8	ABX14570
12	2656.4	72.5	3637	11	ADM68462
13	2656.4	72.5	3637	12	ADP26615
14	2656.4	72.5	3637	13	ADQ88437
15	2656.4	72.5	3637	13	ADQ88437
16	2647.8	72.3	3637	10	ADQ13578
17	2635.8	72.0	3661	2	AAQ13578
18	2622.6	71.6	3314	2	AAQ13576
19	2620.6	71.6	3699	2	AAV14340
20	2609.4	71.3	4001	6	AAD27066

C	21	2517.8	68.8	3918	13	ADR70323	Adr70323 Single ch
C	22	2517.8	68.8	3918	13	ADS17527	Ads17527 Nucleotid
	23	2484.6	67.8	3774	6	AAD27062	Aad27062 Plasmid p
	24	2474.2	67.6	5277	3	AAA88110	Aaa88110 Plasmid p
	25	2464.6	67.3	3908	6	ABT08165	Abt08165 Recombina
	26	2451.8	67.0	3928	8	ABT14478	Abt14478 HCV envel
	27	2451.8	67.0	3928	8	AAD50628	Aad50628 pBSK-E26H
	28	2451.8	67.0	3928	10	ADQ84236	Adq84236 Vector pB
	29	2448.8	66.9	3956	2	AAV64258	Aav64258 Plasmid p
	30	2446.8	66.8	3715	6	AAD27064	Aad27064 Plasmid T
	31	2446.8	66.8	3927	6	ABT08166	Abt08166 Recombina
	32	2314.8	63.2	4088	2	AAV64255	Aav64255 Plasmid p
	33	2305.2	62.9	4102	2	AAV64257	Aav64257 Plasmid p
	34	2279.2	62.2	4119	6	AAD40772	Aad40772 Plasmid p
	35	2278	62.2	3143	13	ADQ88298	Adq88298 High copy
	36	2271.2	62.0	4059	12	ADN97131	Adn97131 Mammalian
	37	2230.6	60.9	5250	10	AAD49957	Aad49957 TAR (Tran
	38	2224.6	60.7	2958	2	AAZ22250	Aaz22250 Nucleotid
	39	2216.2	60.5	4205	4	AAD09088	Aad09088 PERG658 p
	40	2216.2	60.5	4205	6	AAK99713	Aak99713 DNA of pl
C	41	2215.4	60.5	4172	2	AAK19901	Aax19901 Plasmid p
	42	2213	60.4	4205	4	AAD09087	Ada09087 PERG657 p
	43	2213	60.4	4205	6	AAK99700	Aak99700 DNA of pl
	44	2202.6	60.1	4454	4	AAQ06386	Aas06386 Vector PG
	45	2172.8	59.3	7102	2	AAK31836	Aax31836 Nucleotid

ALIGNMENTS

RESULT 1
ACA55354
ID ACA55354 standard; DNA; 3661 BP.
XX
AC ACA55354;
XX
DT 06-JUN-2003 (first entry)
XX
DE Transformation vector piggyBAC related plasmid pXL-Bac.
XX
KW PiggyBac; transposon; eukaryotic transformation vector; ds;
KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;
KW circular.
XX
OS Synthetic.
XX
PN US2002173634-A1.
XX
PD '21-NOV-2002.
XX
PF 30-OCT-2001; 2001US-00001189.
XX
PR 31-OCT-2000; 2000US-0244677P.
XX
PA 01-NOV-2000; 2000US-0244984P.
XX
PA (FRAS/) FRASER M J.
XX
PA (LIXX/) LI X.
XX
PA (BEAM/) BEAM T.
XX
XX (HUA/) HUA-VAN A.
XX
PI Fraser MJ, Li X, Beam T, Hua-Van A;
XX
DR WPI; 2003-352597/33.
XX
PT New DNA molecule in the transposon piggyBac, useful for transferring
PT genes into host cells or embryos for transforming the cells of embryos
PT that can be used in making transgenic organisms.
XX
PS Example 3; Fig 3(C2); 151pp; English.
XX
CC The invention describes a DNA molecule comprising at least 163
CC consecutive nucleotide base pairs of the 3' terminal region beginning at
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base

CC pairs of the 5' terminal region beginning at the 5'terminal base pair, of
CC the piggyBac molecule. The region extends from the restriction site SacI
CC to the end of the piggyBac molecule. The DNA molecule in the transposon
CC piggyBac is useful for transferring genes into host cells or embryos for
CC transforming the cells or embryos. The transformed cells or embryos are
CC useful for developing or making transgenic organisms. This sequence
CC represents plasmid used in the creation of minimal sequence eukaryotic
CC transformation vector piggyBac
XX
SQ Sequence 3661 BP; 910 A; 910 C; 957 G; 884 T; 0 U; 0 Other;

Query Match 99.6%; Score 3648.4; DB 10; Length 3661;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3660; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTAATTTGACGGTTAAATTTTCTGTAATTTTCGATTTTCTGTTAAATCAGCTC 60
DB 1 CTAATTTGACGGTTAAATTTTCTGTAATTTTCGATTTTCTGTTAAATCAGCTC 60
QY 61 ATTTTAAACCAATAGGCGGAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 120
DB 61 ATTTTAAACCAATAGGCGG-ATTCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 119
QY 121 GATAGGGTTAGTGTGTTTCAGTTTGGAAACAAGAGTCCACTATTAAAGAAAGTGGACTC 180
DB 120 GATAGGGTTAGTGTGTTTCAGTTTGGAAACAAGAGTCCACTATTAAAGAAAGTGGACTC 179
QY 181 CAACTGCAAGGGCGGAAACCGTCTATCAGGGCGATGCGCCACTACGTGAACCAATCACC 240
DB 180 CAACTGCAAGGGCGGAAACCGTCTATCAGGGCGATGCGCCACTACGTGAACCAATCACC 239
QY 241 CTAATCAAGTTTTCGGGTCGAGGTGCGTAAAGCACTTAATTCGGAACCCCTAAAGGGAG 300
DB 240 CTAATCAAGTTTTCGGGTCGAGGTGCGTAAAGCACTTAATTCGGAACCCCTAAAGGGAG 299
QY 301 CCCCCGATTTAGAGCTTGAACGGGGAAGCCGCGGAAACGTTGGCGAGAAAGGAAGAA 360
DB 300 CCCCCGATTTAGAGCTTGAACGGGGAAGCCGCGGAAACGTTGGCGAGAAAGGAAGAA 359
QY 361 AGCGAAGAGCGGGCGCTTAGGGCGCTGGCAAGTGTAGGGTACGCTGCGGTAAACAC 420
DB 360 AGCGAAGAGCGGGCGCTTAGGGCGCTGGCAAGTGTAGGGTACGCTGCGGTAAACAC 419
QY 421 CACACCGCGCGCTTAATTCGCGCGCTACAGGGCGCTCCCAATTCGCCATTCAGGCTGCG 480
DB 420 CACACCGCGCGCTTAATTCGCGCGCTACAGGGCGCTCCCAATTCGCCATTCAGGCTGCG 479
QY 481 CAACTGTTGGGAAGGGCGATTCGGTGGGCGCTCTTTCGCTATTACGCCAGCTGGCGAAAGG 540
DB 480 CAACTGTTGGGAAGGGCGATTCGGTGGGCGCTCTTTCGCTATTACGCCAGCTGGCGAAAGG 539
QY 541 GGGATGTCGCAAGCGGATTAAGTTGGGTAAACGCGAGGCTTTCCCACTCAGACGTTG 600
DB 540 GGGATGTCGCAAGCGGATTAAGTTGGGTAAACGCGAGGCTTTCCCACTCAGACGTTG 599
QY 601 TAAACAGCGCGCAGTGAAGCGGCTCTGTTCACTTCACTTCACTTGAACCCGTTGAGAGGACGG 660
DB 600 TAAACAGCGCGCAGTGAAGCGGCTCTGTTCACTTCACTTGAACCCGTTGAGAGGACGG 659
QY 661 GCAGACTCGCGTGCAAAATGTGTTTACAGCGTGATGAGGACAGATGAAGATGCTCGACAC 720
DB 660 GCAGACTCGCGTGCAAAATGTGTTTACAGCGTGATGAGGACAGATGAAGATGCTCGACAC 719
QY 721 GCTGCAGACACGCGAGCTAGATTACCTTAGAAGATTAATCATATTGTGACGTAGTTAA 780
DB 720 GCTGCAGACACGCGAGCTAGATTACCTTAGAAGATTAATCATATTGTGACGTAGTTAA 779
QY 781 AGATAATCATGCTGTAATTTGACGCGATGGGATCTCTTAATACGACTCACTATAGGCGGAAT 840
DB 780 AGATAATCATGCTGTAATTTGACGCGATGGGATCTCTTAATACGACTCACTATAGGCGGAAT 839
QY 841 TGGGTACCGGGCCCCCTTCGAGGTTCGACGGTATCGATAAGCTTGTATTCGAATTCCTGC 900
DB 841 TGGGTACCGGGCCCCCTTCGAGGTTCGACGGTATCGATAAGCTTGTATTCGAATTCCTGC 899

DB 840 TGGGTACCGGGCCCCCTTCGAGGTTCGACGGTATCGATAAGCTTGTATTCGAATTCCTGC 899
QY 901 AGCCCCGGGGATCCACTAGTTCTAGAGCGCGCCGACCGCGTGGAGCTCCAGCTTTTCT 960
DB 900 AGCCCCGGGGATCCACTAGTTCTAGAGCGCGCCGACCGCGTGGAGCTCCAGCTTTTCT 959
QY 961 TCCCTTTAGTAGGGGTTAAATTTAGATCCCATCGTCAATTTTACGAGAGCTATCTTTCTAG 1020
DB 960 TCCCTTTAGTAGGGGTTAAATTTAGATCCCATCGTCAATTTTACGAGAGCTATCTTTCTAG 1019
QY 1021 GGTAAATAGCTGATCAGGATCATATCTGTCGGGTCTTTTTCGGGTCACTATCGCC 1080
DB 1020 GGTAAATAGCTGATCAGGATCATATCTGTCGGGTCTTTTTCGGGTCACTATCGCC 1079
QY 1081 CAAAGCTGCGCTATCTGGGCATCGGGAGGAAGAACCGTCCCTTTTCCCGCAGGTTG 1140
DB 1080 CAAAGCTGCGCTATCTGGGCATCGGGAGGAAGAACCGTCCCTTTTCCCGCAGGTTG 1139
QY 1141 AAGCGGCATGGAAGAGATTGTCGAGGATGACTGCTGTCATTGACGTTGAGCGGAAAC 1200
DB 1140 AAGCGGCATGGAAGAGATTGTCGAGGATGACTGCTGTCATTGACGTTGAGCGGAAAC 1199
QY 1201 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCAACGCTTTAACGCTGAACGT 1260
DB 1200 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCAACGCTTTAACGCTGAACGT 1259
QY 1261 TCGTTTCAGCGCACCTGGGATACCACTTCGTCGGGCTTTTCCGGACACAGTTTCCGGATGG 1320
DB 1260 TCGTTTCAGCGCACCTGGGATACCACTTCGTCGGGCTTTTCCGGACACAGTTTCCGGATGG 1319
QY 1321 TCAGCCGGAAGCGCATCAGCAACCCGAAACAATACCGCGACACGCGGAACCTCCGTCGCG 1380
DB 1320 TCAGCCGGAAGCGCATCAGCAACCCGAAACAATACCGCGACACGCGCGGAACCTCCGTCGCG 1379
QY 1381 GTGTGCAATTAATGACAGCGGTGCGGCTGGGATATTAATGTCAGCGAGGACGGGTATC 1440
DB 1380 GTGTGCAATTAATGACAGCGGTGCGGCTGGGATATTAATGTCAGCGAGGACGGGTATC 1439
QY 1441 CTGGCTGATGCGCGAGAAATGGACATGATACCCGCTGAGTTTACCCGCGCGCGCTT 1500
DB 1440 CTGGCTGATGCGCGAGAAATGGACATGATACCCGCTGAGTTTACCCGCGCGCGCTT 1499
QY 1501 GCGGTAATCATGCTCATAGCTGTTTCTGTTGAAATTTTATTCGCTCACAATTTCCACA 1560
DB 1500 GCGGTAATCATGCTCATAGCTGTTTCTGTTGAAATTTTATTCGCTCACAATTTCCACA 1559
QY 1561 CAACATAGCGCGGAGCAATAAAGTGTAAAGCTTGGGGTGCCTAATGAGTGAGCTAAT 1620
DB 1560 CAACATAGCGCGGAGCAATAAAGTGTAAAGCTTGGGGTGCCTAATGAGTGAGCTAAT 1619
QY 1621 CACATTAATTTGCTTTCGCTCACTGCGCGCTTTCCAGTTCGGGAAACCTGTCTGTCAGCT 1680
DB 1620 CACATTAATTTGCTTTCGCTCACTGCGCGCTTTCCAGTTCGGGAAACCTGTCTGTCAGCT 1679
QY 1681 GCATTAATGAATTCGCGCAACCGCGGAGAGCGGTTTGGTATTTGGGCGCTCTTCGCG 1740
DB 1680 GCATTAATGAATTCGCGCAACCGCGGAGAGCGGTTTGGTATTTGGGCGCTCTTCGCG 1739
QY 1741 TTCTCTGCTCACTGATCTGCTGCGCTCGGTTCGCTGCGGCGAGCGGTATGAGCTCA 1800
DB 1740 TTCTCTGCTCACTGATCTGCTGCGCTCGGTTCGCTGCGGCGAGCGGTATGAGCTCA 1799
QY 1801 CTCAGAGCGGTAAATACGTTTATCCACAGATTAACGAGGATTAACGAGGAAACATGTG 1860
DB 1800 CTCAGAGCGGTAAATACGTTTATCCACAGATTAACGAGGATTAACGAGGAAACATGTG 1859
QY 1861 AGCAAAAGCCAGCAAAAGGCGGAAACCGTAAAGCCCGGCTTGTGCGGCTTTTTC 1920
DB 1860 AGCAAAAGCCAGCAAAAGGCGGAAACCGTAAAGCCCGGCTTGTGCGGCTTTTTC 1919
QY 1921 TAGGCTCCGCGCCCCCTTGACGAGCATCAAAAATTCGACGCTCAAGTCAAGAGGTGCGGAA 1980
DB 1920 TAGGCTCCGCGCCCCCTTGACGAGCATCAAAAATTCGACGCTCAAGTCAAGAGGTGCGGAA 1979

1981 CCGGACAGGACTATAAAGATACACAGGCGTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCC 2040
1980 CCGGACAGGACTATAAAGATACACAGGCGTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCC 2039
2041 TGTTCGACCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTCCGGAAGCGTGGC 2100
2040 TGTTCGACCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTCCGGAAGCGTGGC 2099
2101 GCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGGTTCGCTCCAAGCT 2160
2100 GCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGGTTCGCTCCAAGCT 2159
2161 GGGCTGTGTGACGAACCCCTGTTAGCGGCGACCGGTGGCTTATCCGGTAACTATCG 2220
2160 GGGCTGTGTGACGAACCCCTGTTAGCGGCGACCGGTGGCTTATCCGGTAACTATCG 2219
2221 TCTTGAGTCCAAACCCGTAAGACACACTTATCGCACTGTCGACGAGCCACTGGTAACAG 2280
2220 TCTTGAGTCCAAACCCGTAAGACACACTTATCGCACTGTCGACGAGCCACTGGTAACAG 2279
2281 GATTACGACGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTAACTA 2340
2280 GATTACGACGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTAACTA 2339
2341 CGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGTAAGCCAGTTACCTTCGG 2400
2340 CGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGTAAGCCAGTTACCTTCGG 2399
2401 AAAAAAGAGTTGGTAGTCTTGTATCGGCAAAACAAACACCGCTGGTAGCGGTGTTTTT 2460
2400 AAAAAAGAGTTGGTAGTCTTGTATCGGCAAAACAAACACCGCTGGTAGCGGTGTTTTT 2459
2461 TGTTCGACAGCAGATACGCGCAGAAAAAAGGATCTCAAGAGATGCTTTTGATCTT 2520
2460 TGTTCGACAGCAGATACGCGCAGAAAAAAGGATCTCAAGAGATGCTTTTGATCTT 2519
2521 TTCTACGGGCTGACGCTCAGTGGAAACGAAAACTCACGCTTAAAGGATTTGGTCAATGAG 2580
2520 TTCTACGGGCTGACGCTCAGTGGAAACGAAAACTCACGCTTAAAGGATTTGGTCAATGAG 2579
2581 ATTATCAAAAAGGATCTTACCTAGATCTTTTAAATTAAGATGAATTTAAATCAAT 2640
2580 ATTATCAAAAAGGATCTTACCTAGATCTTTTAAATTAAGATGAATTTAAATCAAT 2639
2641 CTAAAGTATATAGTAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2700
2640 CTAAAGTATATAGTAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2699
2701 TATCTCAGCGATCTGTCTATTTTCGTTTATCCATAGTTGCTGACTCCCGCTCGTGTAGAT 2760
2700 TATCTCAGCGATCTGTCTATTTTCGTTTATCCATAGTTGCTGACTCCCGCTCGTGTAGAT 2759
2761 AACTAGATACGGAGGGCTTACCATCTGGCCCGAGTGTGCAATGATATACCGGACACC 2820
2760 AACTAGATACGGAGGGCTTACCATCTGGCCCGAGTGTGCAATGATATACCGGACACC 2819
2821 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACACGACGCGGAGGCGGCGAG 2880
2820 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACACGACGCGGAGGCGGCGAG 2879
2881 AAGTGTCTCTGCAACTTTATCCGCTCCATCAGTCTATTAATTTGTCGGGAAGCTAG 2940
2880 AAGTGTCTCTGCAACTTTATCCGCTCCATCAGTCTATTAATTTGTCGGGAAGCTAG 2939
2941 AGTAAGTAGTTCGCCAGTTAAATAGTTTGGCGAAGCGTTGTCATGCTACAGGCAATCGT 3000
2940 AGTAAGTAGTTCGCCAGTTAAATAGTTTGGCGAAGCGTTGTCATGCTACAGGCAATCGT 2999
3001 GGTGTCAAGCTCGTCTGTTGGTATGCTTCAATTCAGCTCGGTTCCCAACGATCAAGGCG 3060
3000 GGTGTCAAGCTCGTCTGTTGGTATGCTTCAATTCAGCTCGGTTCCCAACGATCAAGGCG 3059

3061 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCTTTCGGTCTCCGATCGT 3120
3060 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCTTTCGGTCTCCGATCGT 3119
3121 TGTCAAGAGTAAGTTGCGCGCAGTGTATACACTCATGGTTATGCGCAGCACTGCATAATTC 3180
3120 TGTCAAGAGTAAGTTGCGCGCAGTGTATACACTCATGGTTATGCGCAGCACTGCATAATTC 3179
3181 TCTTACTGTCAATCCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTC 3240
3180 TCTTACTGTCAATCCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTC 3239
3241 ATTCTGAGATAGTGTATGCGGCGACCGAGTGTCTTTCGCCGCGCTCAATACGGATAA 3300
3240 ATTCTGAGATAGTGTATGCGGCGACCGAGTGTCTTTCGCCGCGCTCAATACGGATAA 3299
3301 TACCGCGCACATAGCAGAACTTTAAAAAGTGTCTCATATTGGAACGTTCTTTCGGGCG 3360
3300 TACCGCGCACATAGCAGAACTTTAAAAAGTGTCTCATATTGGAACGTTCTTTCGGGCG 3359
3361 AAAAATCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCAATGTAACCCACTCGTGACC 3420
3360 AAAAATCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCAATGTAACCCACTCGTGACC 3419
3421 CAACTGATCTCAGCATCTTTTACTTTTACAGCGGTTTCTGGGTGAGCAAAAAACAGGAAG 3480
3420 CAACTGATCTCAGCATCTTTTACTTTTACAGCGGTTTCTGGGTGAGCAAAAAACAGGAAG 3479
3481 GCAAAATGCGCAAAAAAGGAATAGGCGCACCGGAAATGTTGAATGTTGAATCTCATCTT 3540
3480 GCAAAATGCGCAAAAAAGGAATAGGCGCACCGGAAATGTTGAATGTTGAATCTCATCTT 3539
3541 CTTTTTCAATATATTATGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
3540 CTTTTTCAATATATTATGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3599
3601 TGAATGTTATGAAAAATAAACAATAGGGGTTCCGCGCATTTTCCCGAAAAAGTGCC 3660
3600 TGAATGTTATGAAAAATAAACAATAGGGGTTCCGCGCATTTTCCCGAAAAAGTGCC 3659
3661 AC 3662
3660 AC 3661

RESULT 2
ADM68463/C
ID ADM68463 standard; DNA; 3637 BP.
XX
AC ADM68463;
XX
DT 03-JUN-2004 (first entry)
XX
DE Variant green fluorescent protein, GFP, gene construct.
XX
KW ds; mismatch endonuclease; endonuclease; gene shuffling technology;
KW single nucleotide polymorphism; cancer susceptibility;
KW sequence variation redistribution; GFP; green fluorescent protein; gene.
XX
OS Aequorea victoria.
OS Synthetic.
XX
PN US2003157682-A1.
XX
PD 21-AUG-2003.
XX
PF 31-JAN-2003; 2003US-00356708.
XX
PR 01-FEB-2002; 2002US-0353722P.
PR 14-MAR-2002; 2002US-00098155.
PR 01-AUG-2002; 2002US-00211079.
XX
PA (PADG/) PADGETT H S.

QY 1620 TCACATTAATTGCGTTGCGCTCACTGCCCCCTTTCCAGTCGGAAACCTGTCGTGCGAGC 1679
DB 2043 TCACATTAATTGCGTTGCGCTCACTGCCCCCTTTCCAGTCGGAAACCTGTCGTGCGAGC 1984
QY 1680 TGCATTAATCAATCGGCCCAACGCGCGGAGAGCGGTTGCGTATTTGGCGCTCTTCCG 1739
DB 1983 TGCATTAATGAATTCGGCCCAACGCGCGGAGAGCGGTTTTCGCTATTTGGCGCTCTTCCG 1924
QY 1740 CTTTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGCGAGCGGTATCAGCTC 1799
DB 1923 CTTTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGCGAGCGGTATCAGCTC 1864
QY 1800 ACTCAAAAGCGGTAAATACGGTTATCCACAGAATCAGGGGATACCGACGAAAGAAATGT 1859
DB 1863 ACTCAAAAGCGGTAAATACGGTTATCCACAGAATCAGGGGATACCGACGAAAGAAATGT 1804
QY 1860 GAGCAAAAGCCAGCAAAAGCCAGAAACCGTAAAGGCGGTTGCTGGCGCTTTTTC 1919
DB 1803 GAGCAAAAGCCAGCAAAAGCCAGAAACCGTAAAGGCGGTTGCTGGCGCTTTTTC 1744
QY 1920 ATAGGCTCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAA 1979
DB 1743 ATAGGCTCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAA 1684
QY 1980 ACCGACAGGACTATTAAGATACAGCGGTTTCCCTCGTGAAGCTCCCTCGTGCCTCTC 2039
DB 1683 ACCGACAGGACTATTAAGATACAGCGGTTTCCCTCGTGAAGCTCCCTCGTGCCTCTC 1624
QY 2040 CTGTTCCGACCTCGCGCTACCGGATACCTGTCGCGCTTTCCTTCGGGAGCGTGG 2099
DB 1623 CTGTTCCGACCTCGCGCTACCGGATACCTGTCGCGCTTTCCTTCGGGAGCGTGG 1564
QY 2100 CGCTTTCTCATAGCTCACGCTGTAGTATCTCAGTTCCGTTGAGTGTGCTTCCGCTCAAGC 2159
DB 1563 CGCTTTCTCATAGCTCACGCTGTAGTATCTCAGTTCCGTTGAGTGTGCTTCCGCTCAAGC 1504
QY 2160 TGGGCTGTGTGACGAAACCCCGTTTCAGCCGACCGCTGCGCTTATCCGTTAACTATC 2219
DB 1503 TGGGCTGTGTGACGAAACCCCGTTTCAGCCGACCGCTGCGCTTATCCGTTAACTATC 1444
QY 2220 GTCTTGAGTCCAAACCGGTAAAGCAGCACTATCCGCACTGGCAGCAGCACTGGTAAAC 2279
DB 1443 GTCTTGAGTCCAAACCGGTAAAGCAGCACTATCCGCACTGGCAGCAGCACTGGTAAAC 1384
QY 2280 GGATTAAGCAGGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCTTAATC 2339
DB 1383 GGATTAAGCAGGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCTTAATC 1324
QY 2340 ACGGCTTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG 2399
DB 1323 ACGGCTTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG 1264
QY 2400 GAAAAGAGTTGGTAGCTCTTGATCCGCGCAACCAACCCGCTGCTAGCGGTGGTTTTT 2459
DB 1263 GAAAAGAGTTGGTAGCTCTTGATCCGCGCAACCAACCCGCTGCTAGCGGTGGTTTTT 1204
QY 2460 TTGTTTGAAGCAGCAGATTAAGCGCAGAAAAAAGGATCTCAAGAGATCTCTTGAATCT 2519
DB 1203 TTGTTTGAAGCAGCAGATTAAGCGCAGAAAAAAGGATCTCAAGAGATCTCTTGAATCT 1144
QY 2520 TTTCTACGGGCTGTGACGCTCAGTGGAACGAAACCTCAGGTTAAGGATTTTGGTCATGA 2579
DB 1143 TTTCTACGGGCTGTGACGCTCAGTGGAACGAAACCTCAGGTTAAGGATTTTGGTCATGA 1084
QY 2580 GATTATCAAAAGGATCTTCACTAGATCTCTTTTAAATTTAAAAATCAAGTTTAAATCAA 2639
DB 1083 GATTATCAAAAGGATCTTCACTAGATCTCTTTTAAATTTAAAAATCAAGTTTAAATCAA 1024
QY 2640 TCTAAAGTATATATGATAGTAAACTTGGTCTGACAGTTTACCAATGCTTTAATCAGTAGGAC 2699
DB 1023 TCTAAAGTATATATGATAGTAAACTTGGTCTGACAGTTTACCAATGCTTTAATCAGTAGGAC 964

QY 2700 CTATCTCAGGAGTCTCTCTATTTCTGTTTCATTCATAGTTGCTGCTGACTCCCGCTGTTAGA 2759
DB 963 CTATCTCAGGAGTCTCTCTATTTCTGTTTCATTCATAGTTGCTGCTGACTCCCGCTGTTAGA 904
QY 2760 TAACTACGATACGGGAGGCTTACCATCTGGCCCCAGTCTGCAATGATATACCGGAGACC 2819
DB 903 TAACTACGATACGGGAGGCTTACCATCTGGCCCCAGTCTGCAATGATATACCGGAGACC 844
QY 2820 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCAGCGGAAAGCCGAGCGCA 2879
DB 843 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCAGCGGAAAGCCGAGCGCA 784
QY 2880 GAAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCGGGAAGCTA 2939
DB 783 GAAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCGGGAAGCTA 724
QY 2940 GAGTAAGTAGTTTCGCGCAGTTAATAGTTTGCACAACTGTTGTCATTTGCTACAGGATCG 2999
DB 723 GAGTAAGTAGTTTCGCGCAGTTAATAGTTTGCACAACTGTTGTCATTTGCTACAGGATCG 664
QY 3000 TGGTGTACGCTCGTCTGTTTGGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGGC 3059
DB 663 TGGTGTACGCTCGTCTGTTTGGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGGC 604
QY 3060 GAGTTACATGATCCCGCATGTTGTGCAAAAAGCGGTAGCTCCTTCGCTCCGATCG 3119
DB 603 GAGTTACATGATCCCGCATGTTGTGCAAAAAGCGGTAGCTCCTTCGCTCCGATCG 544
QY 3120 TTGTCAAGAGTAAAGTTGGCGCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATT 3179
DB 543 TTGTCAAGAGTAAAGTTGGCGCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATT 484
QY 3180 CTCTTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAATCTCAACCAAGT 3239
DB 483 CTCTTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAATCTCAACCAAGT 424
QY 3240 CAATCTGAAGATAGTGTATGCGCGGACCGAGTTGCTCTTCCCGCGGTCAATACGGGATA 3299
DB 423 CAATCTGAAGATAGTGTATGCGCGGACCGAGTTGCTCTTCCCGCGGTCAATACGGGATA 364
QY 3300 ATACCGCGCACATAGCAGAACTTTTAAAGTGTCTCATCTGGAATAAGTTCTTCGGGGC 3359
DB 363 ATACCGCGCACATAGCAGAACTTTTAAAGTGTCTCATCTGGAATAAGTTCTTCGGGGC 304
QY 3360 GAAAACTCTCAAGGATCTTACCGCTGTTCAGATGCCAGTTCCAGTGAACCCACCTCGTGCAC 3419
DB 303 GAAAACTCTCAAGGATCTTACCGCTGTTCAGATGCCAGTTCCAGTGAACCCACCTCGTGCAC 244
QY 3420 CCAACTGTATCTTCAGCATCTTTTACCTTTTCAAGCGTTTCTGGGTGAGCAAAAACAGGAA 3479
DB 243 CCAACTGTATCTTCAGCATCTTTTACCTTTTCAAGCGTTTCTGGGTGAGCAAAAACAGGAA 184
QY 3480 GSCAAAATGSCCAAAAAGGGAATAAGGCGCAGACGGAATAAGTTCGAATCTCATACTCT 3539
DB 183 GSCAAAATGSCCAAAAAGGGAATAAGGCGCAGACGGAATAAGTTCGAATCTCATACTCT 124
QY 3540 TCCTTTTCAATATTTAAGCACTTTATCAGGTTTATTTGCTCTCATGAGCGGATACATAT 3599
DB 123 TCCTTTTCAATATTTAAGCACTTTATCAGGTTTATTTGCTCTCATGAGCGGATACATAT 64
QY 3600 TTGAATGTATTTAGAAAAATAAATAAGGGGTTCCCGCGCAGCATTTCCCGAAAAAGTGC 3659
DB 63 TTGAATGTATTTAGAAAAATAAATAAGGGGTTCCCGCGCAGCATTTCCCGAAAAAGTGC 4
QY 3660 CAC 3662
DB 3 CAC 1

RESULT 3
ADP26616/c
ID ADP26616 standard; DNA; 3637 BP.
XX

QY 1260 TTGGTTGAGGACCTGGGATACAGTTCGTGGGGCTTTTCGGGACACAGTTCGGATG 1319
Db 2382 TGCCCATGGAACAGGTAGTTTTCAGTAGTGCAATAAATTTAAGTGTAAAGCTTTCCGTA 2323
QY 1320 GTGAGCCGAAGCGCATCAGCAACCCGAAACAATACCGGGGACAGCCGGAATGCGGTGC 1379
Db 2322 TGATGATCACCTTACCTCTCCACTGACAGAAATTTGTGCCATTAACATCACCATC 2263
QY 1380 GGTGTGCAGATTAATGACAGCGGTGGCGGTGGGATATTAAGTACGAGGAGCGGTAT 1439
Db 2262 TAAATCAACAAGAAAT-----TGGGACAACTCCAGTGAAAAAGTTC 2224
QY 1440 CTTGGCTGATCCGACAGAAATGACATGATGATACCCCGTGAATACCCGGCGGCGCT 1499
Db 2223 TTCCTCTTACTCATCGTACCAGCTTTTGTTCCTTTAGTGAGGGTTAAATTTGGCGCT 2164
QY 1500 TGGCGTAATCATGCTCATAGCTGTTCCTGTGTGAATTTGTTATCCGCTCACAAATCCAC 1559
Db 2163 TGGCGTAATCATGCTCATAGCTGTTCCTGTGTGAATTTGTTATCCGCTCACAAATCCAC 2104
QY 1560 ACAACATACGAGCCGGAAGATAAAGTAAAGCCTGGGGTGCCTAATGAGTGAGTAAAC 1619
Db 2103 ACAACATACGAGCCGGAAGATAAAGTAAAGCCTGGGGTGCCTAATGAGTGAGTAAAC 2044
QY 1620 TCACATTAATTTGGTTCGCTCACTGCGCTCACTGCGCTTTCAGTCCGGAACCTGTCGTCAGC 1679
Db 2043 TCACATTAATTTGGTTCGCTCACTGCGCTTTCAGTCCGGAACCTGTCGTCAGC 1984
QY 1680 TGCAATTAATGAAATCGGCAAGCGGGGAGAGCGGTTTGGCTATTTGGCGCTCTTCCG 1739
Db 1983 TGCAATTAATGAAATCGGCAAGCGGGGAGAGCGGTTTGGCTATTTGGCGCTCTTCCG 1924
QY 1740 CTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1799
Db 1923 CTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1864
QY 1800 ACTCAAGCGGTATACGGTTATCACAAGATCAGGGATACGCGGAAGAACATGT 1859
Db 1863 ACTCAAGCGGTATACGGTTATCACAAGATCAGGGATACGCGGAAGAACATGT 1804
QY 1860 GAGCAAAAGCGCAGCAAAAGCGCAGAAACCGTAAAGGCGGTTGCTGGCGTTTTCG 1919
Db 1803 GAGCAAAAGCGCAGCAAAAGCGCAGAAACCGTAAAGGCGGTTGCTGGCGTTTTCG 1744
QY 1920 ATAGGCTCCGCCCTCTGACGAGCATCACAAAAATCAGCTCAAGTCAAGGTGCGGAA 1979
Db 1743 ATAGGCTCCGCCCTCTGACGAGCATCACAAAAATCAGCTCAAGTCAAGGTGCGGAA 1684
QY 1980 ACCGACGAGCATTAAGATACGAGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTC 2039
Db 1683 ACCGACGAGCATTAAGATACGAGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTC 1624
QY 2040 CTGTTCCGACCTGCGCTTACCGGATCTGTGCGCTTCTCCCTCGGGAAGCGTG 2099
Db 1623 CTGTTCCGACCTGCGCTTACCGGATCTGTGCGCTTCTCCCTCGGGAAGCGTG 1564
QY 2100 CGCTTTCTCATAGCTCACGCTGTAGTATCTCAGTTTCGCTGTAGTTCGCTCCAAGC 2159
Db 1563 CGCTTTCTCATAGCTCACGCTGTAGTATCTCAGTTTCGCTGTAGTTCGCTCCAAGC 1504
QY 2160 TGGGTGTGTGACGAAACCCCGTTACGCGGACCGCTGGCGCTTATCCGGTAATATC 2219
Db 1503 TGGGTGTGTGACGAAACCCCGTTACGCGGACCGCTGGCGCTTATCCGGTAATATC 1444
QY 2220 GTCTTGAGTCCAAACCGGTAAAGACGACTTATCGCACTGCGAGCAGCACTGGTAACA 2279
Db 1443 GTCTTGAGTCCAAACCGGTAAAGACGACTTATCGCACTGCGAGCAGCACTGGTAACA 1384
QY 2280 GGAATAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTTGAAGTGGTGGCTTAAT 2339
Db 1383 GGAATAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTTGAAGTGGTGGCTTAAT 1324
QY 2340 ACGGCTACACTAGAGGACAGTATTTGGTATCTGGCGCTCTGCTGAAGCCAGCTTACCTTCG 2399

Db 1323 ACGGCTACACTAGAGGACAGTATTTGGTATCTGCCTCTGCTGAAGCCAGTTACCTTCG 1264
QY 2400 GAAAAAGAGTTGGTAGCTCTTGATCCGCAAAACAAACCCAGCTGGTAGCGGTGTTTT 2459
Db 1263 GAAAAAGAGTTGGTAGCTCTTGATCCGCAAAACAAACCCAGCTGGTAGCGGTGTTTT 1204
QY 2460 TTGTTTGAAGCAGAGATACCGGAGAAAAAAGGATCTCAAGAGATCTTTGATCT 2519
Db 1203 TTGTTTGAAGCAGAGATACCGGAGAAAAAAGGATCTCAAGAGATCTTTGATCT 1144
QY 2520 TTTCTACGGGTCTGACGCTCAGTGGAACGAAACCTCAGTTAAGGATTTTGGTCATGA 2579
Db 1143 TTTCTACGGGTCTGACGCTCAGTGGAACGAAACCTCAGTTAAGGATTTTGGTCATGA 1084
QY 2580 GATTATCAAAAAGGATCTTACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAA 2639
Db 1083 GATTATCAAAAAGGATCTTACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAA 1024
QY 2640 TCTAAAGTATATATGAGTAAACTTTGGTCTGACAGATTAACAATCTTAAATCAGTGAGCAC 2699
Db 1023 TCTAAAGTATATATGAGTAAACTTTGGTCTGACAGATTAACAATCTTAAATCAGTGAGCAC 964
QY 2700 CTATCTCAGGATCTGTCTATTTTGGTTCATATGCTGACTCCCGCTCGTGTAGA 2759
Db 963 CTATCTCAGGATCTGTCTATTTTGGTTCATATGCTGACTCCCGCTCGTGTAGA 904
QY 2760 TAACTACGATACGGGAGGCTTACCATCTGGGCCAGTGTGCAATGATACCCGAGACC 2819
Db 903 TAACTACGATACGGGAGGCTTACCATCTGGGCCAGTGTGCAATGATACCCGAGACC 844
QY 2820 CACGCTCACCGGCTCAGATTTATCAGCAATAAACCAGCAGCGGAGGCGCAGCGCA 2879
Db 843 CACGCTCACCGGCTCAGATTTATCAGCAATAAACCAGCAGCGGAGGCGCAGCGCA 784
QY 2880 GAAGTGTCTCCCACTTTATCGCTCCATCCAGTCTAATTAATCTGCGCGGAGCTA 2939
Db 783 GAAGTGTCTCCCACTTTATCGCTCCATCCAGTCTAATTAATCTGCGCGGAGCTA 724
QY 2940 GAGTAAGTAGTTTCGGCAGTAAATAGTTTTCGCAACCGTTGTGCAATGCTACAGGATCG 2999
Db 723 GAGTAAGTAGTTTCGGCAGTAAATAGTTTTCGCAACCGTTGTGCAATGCTACAGGATCG 664
QY 3000 TGGTGTACGCTCGCTGTTTGGTATGGCTTCAATCAGCTCCGTTCCCAACGATCAAGGC 3059
Db 663 TGGTGTACGCTCGCTGTTTGGTATGGCTTCAATCAGCTCCGTTCCCAACGATCAAGGC 604
QY 3060 GAGTTACATGATCCCGATGTTGCAAAAAGCGGTTAGCTCTTCCGCTCCGATCG 3119
Db 603 GAGTTACATGATCCCGATGTTGCAAAAAGCGGTTAGCTCTTCCGCTCCGATCG 544
QY 3120 TTGTCAGAAAGTAAGTTGGCGCAGTGTATCACTCATGTGTTATGGCAGCACTGCATAAT 3179
Db 543 TTGTCAGAAAGTAAGTTGGCGCAGTGTATCACTCATGTGTTATGGCAGCACTGCATAAT 484
QY 3180 CTCTTACTGTCTAGCCATCCGTAAGATGCTTTTGTGTACTGGTGTACTCAACCAAGT 3239
Db 483 CTCTTACTGTCTAGCCATCCGTAAGATGCTTTTGTGTACTGGTGTACTCAACCAAGT 424
QY 3240 CATTCGAGAAATAGTGTATGCGGACCGAGTTGCTTTCGCCGGGCTCAATACGGGATA 3299
Db 423 CATTCGAGAAATAGTGTATGCGGACCGAGTTGCTTTCGCCGGGCTCAATACGGGATA 364
QY 3300 ATACCGGCGCACATAGCAAACTTTTAAAGTGTCTCATCATGTAAGAAAGCTTTTCGGGGC 3359
Db 363 ATACCGGCGCACATAGCAAACTTTTAAAGTGTCTCATCATGTAAGAAAGCTTTTCGGGGC 304
QY 3360 GAAAACTCTCAAGGATCTTACCGCTGTTCAGATCCAGTTTCGATGTAAACCCACTCGTGAC 3419
Db 303 GAAAACTCTCAAGGATCTTACCGCTGTTCAGATCCAGTTTCGATGTAAACCCACTCGTGAC 244
QY 3420 CCAACTGATCTTCAGGATCTTTTACTTTTCCAGCGTTTCTGGGTGAGCAAAAACAGGAA 3479

Db 243 CCAACTGATCTTTCAGCATCTTTACTTACCTTCCACGAGGTTTCTGGTGCAGCAAAACAGGAA 184
 QY 3480 GGCAGAAATGCCGCAAAAGGGAATAAGGGCCACACGGAATGTTGAATACATCATCTCT 3539
 Db 183 GGCAGAAATGCCGCAAAAGGGAATAAGGGCCACACGGAATGTTGAATACATCATCTCT 124
 QY 3540 TCCTTTTCAATATATTGAAGCATTTATTCAGGGTTATTTGCTCATGAGCGGATACATAT 3599
 Db 123 TCCTTTTCAATATATTGAAGCATTTATTCAGGGTTATTTGCTCATGAGCGGATACATAT 64
 QY 3600 TTGAATGATTTAGAAAAATAAACAATAGGGTTCCGGGCAATTTCCCGGAAAGTGC 3659
 Db 63 TTGAATGATTTAGAAAAATAAACAATAGGGTTCCGGGCAATTTCCCGGAAAGTGC 4
 QY 3660 CAC 3662
 Db 3 CAC 1

RESULT 4

ADQ88438/c
 ID ADQ88438 standard; DNA; 3637 BP.
 XX
 AC ADQ88438;
 DT 18-NOV-2004 (first entry)
 XX
 DE pBS3BFP DNA encoding Jellyfish cycle 3 BFP protein.
 XX
 KW Molecular biology; jellyfish; blue fluorescent protein; BFP; ds.
 XX
 OS Aequorea victoria.
 XX
 FN US2004142433-A1.
 XX
 PD 22-JUL-2004.
 XX
 PF 10-OCT-2003; 2003US-00684134.
 XX
 PR 02-FEB-2001; 2001US-0266386P.
 PR 14-FEB-2001; 2001US-0268788P.
 PR 01-FEB-2002; 2002US-00066390.
 PR 08-AUG-2002; 2002US-0402342P.
 PR 21-AUG-2002; 2002US-00226372.
 PR 25-OCT-2002; 2002US-00280913.
 PR 08-AUG-2003; 2003US-00637758.
 XX

(PADG/) PADGETT H S.
 PA (FITZ/) FITZMAURICE W P.
 PA (LIND/) LINDBO J A.
 PA (VAEW/) VAETHHONGS A A.
 PA (VOJD/) VOJDANI F S.
 PA (SMIT/) SMITH M L.
 XX
 Padgett HS, Fitzmaurice WP, Lindbo JA, Vaethhongs AA, Vojdani FS,
 PI Smith ML;
 XX
 WPI; 2004-552565/53.
 XX
 Preparing variant polynucleotides having different nucleotide sequences
 PT from at least two parent polynucleotides, useful in molecular biology, in
 PT particular for generating populations of related nucleic acid molecules.
 XX
 Example 14; SEQ ID NO 32; 81pp; English.
 PS
 XX The invention relates to a method of preparing a variant polynucleotide
 CC having a different nucleotide sequence from at least two parent
 CC polynucleotides. The method involves preparing at least one heteroduplex
 CC between the two parent polynucleotides, cleaving at least one
 CC polynucleotide strand in the heteroduplex at a mismatch site to form a
 CC cleavage site, replacing at least one nucleotide on at least one strand
 CC at or near the cleavage site where at least one of the strands has a
 CC different nucleotide sequence from either of the at least two parent

CC polynucleotides. Methods and compositions of the invention are useful in
 CC the field of molecular biology in particular for generating populations
 CC of related nucleic acid molecules. The present sequence is pBS3BFP
 CC plasmid DNA encoding Jellyfish cycle 3 BFP (blue fluorescent protein).
 XX This sequence is used in the exemplification of the invention
 XX
 SQ Sequence 3637 BP; 942 A; 891 C; 870 G; 934 T; 0 U; 0 Other;
 Query Match 72.8%; Score 2665.4; DB 13; Length 3637;
 Best Local Similarity 83.7%; Pred. No. 0;
 Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;
 QY 1 CTAATTCGTAAGCGTTAATATTTTCTTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60
 Db 3637 CTAATTCGTAAGCGTTAATATTTTCTTAAATTCGCGTTAAATTTTGTAAATCAGCTC 3578
 QY 61 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATAAATCAAAAGATAGACCGA 120
 Db 3577 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATAAATCAAAAGATAGACCGA 3518
 QY 121 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAAGCTGACTC 180
 Db 3517 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAAGCTGACTC 3458
 QY 181 CAACGTCAAAGGGCGAAACCCGCTCTATCAGGGCGATGCGCCACTACGTTGAACCATCAC 240
 Db 3457 CAACGTCAAAGGGCGAAACCCGCTCTATCAGGGCGATGCGCCACTACGTTGAACCATCAC 3398
 QY 241 CTAATCAAGTTTGTGGGTGAGTGCGGTAAAGCACTAATCGGAACCCCTAAAGGGAG 300
 Db 3397 CTAATCAAGTTTGTGGGTGAGTGCGGTAAAGCACTAATCGGAACCCCTAAAGGGAG 3338
 QY 301 CCCCAGATTTAGAGCTTCACGCGGAAAGCCGCGCAACCTGGCGAGAAAGGAGGAA 360
 Db 3337 CCCCAGATTTAGAGCTTCACGCGGAAAGCCGCGCAACCTGGCGAGAAAGGAGGAA 3278
 QY 361 AGCGAAAGAGCGGCGCTAGGCGCTGCGCAAGTGTAGCGTACGCTGCGGTAAACAC 420
 Db 3277 AGCGAAAGAGCGGCGCTAGGCGCTGCGCAAGTGTAGCGTACGCTGCGGTAAACAC 3218
 QY 421 CACACCCGCGCTTAATGCGCGCTACAGGGCGGTCCTCCATTCGCCATTCAGGCTGCG 480
 Db 3217 CACACCCGCGCTTAATGCGCGCTACAGGGCGGTCCTCCATTCGCCATTCAGGCTGCG 3158
 QY 481 CAACCTGTTGGGAAGGCGCATCGTGGCGCTCTTCGCTATTACGCCAGTGGCGAAAGG 540
 Db 3157 CAACCTGTTGGGAAGGCGCATCGTGGCGCTCTTCGCTATTACGCCAGTGGCGAAAGG 3098
 QY 541 GGGATGCTGCTCAAGGCGGATTAAGTTGGGTAAACCGAGGGTTTCCAGTCAAGCGTTG 600
 Db 3097 GGGATGCTGCTCAAGGCGGATTAAGTTGGGTAAACCGAGGGTTTCCAGTCAAGCGTTG 3038
 QY 601 TAAACGACGCGCAGTGAAGCGGCTCGTTCACTTACGCTTTTGAACCGGTGAGGACGG 660
 Db 3037 TAAACGACGCGCAGTGAAGCGGCTCGTTCACTTACGCTTTTGAACCGGTGAGGACGG 2978
 QY 661 GCAGACTCGCGTGCAAAATGTTTACAGCGTGTAGGAGAGATGCTCGACAC 720
 Db 2977 CGCGGTGCGCGCTCTAGAACTAGTGGATCCCGCGGCTGCAGGAATCTTATTGTT 2918
 QY 721 GCTGCAGAACGCGCAGCTAGATTAAACCTAGAAAGATATCATATTGTGACGTATGTA 780
 Db 2917 ATAGTTTCATCCATGCTGTGTAAATCCAGCAGCAGTGTACAACTCAAGAGAGGACCATGT 2858
 QY 781 AGATATTC-ATCGGTAAATTTGACCGCATGGGATCGTAATACGACTCACTATAGGCGAA 839
 Db 2857 GGTCAACGCTTTTCGTTGGGATCTTTTCAAGAGGCGCAGATTGTGTCGACAGGTAATGTTGT 2798
 QY 840 TTGGGTACCGGCGCCCGCTCGAGGTGCGAGGTATCGATTAAGCTTGCATATCGAATTCCTG 899
 Db 2797 CTGGTAAAGGACAGGCGCATCGCAATTTGGAGTATTTTGTGTGATAATGGTCTGCTAGTT 2738
 QY 900 CAGCCCGGGGATCCACTAGTTTCTAGAGCGGCGCCACCGCGGTGGAGCTCCAGCTTTTG 959

Db 2737 GAAAGGATCCATCTTCAATGTTGTGGGAATTTTGAAGTTAGCTTTGATTCATCTTTT 2678
Qy 960 TTCCCTTTTGTAGGGTTAAATAGATCCCATGCGTCAATTTTACGAGACTATCTTTCTA 1019
Db 2677 GTTTGTCTGCGGTGATATACATTTGTGAGTTAAAGTTGTACTCGAGTTGTGTCCGA 2618
Qy 1020 GGGTTAACTAGCTGATCAGGATCAATGCTGCGGTCTTTTTCGCGCTCAGTCAATCGC 1079
Db 2617 GAAATGTTTCCATCTTCTTTTAAATCAATACCTTTTAACTCGATAC-----GATTAAACAAG 2563
Qy 1080 CCAAGCTGGCGTATCTGGGCATCGGGAGGAAGACCGTCTTTTCCCGGAGGTT 1139
Db 2562 GGTATCACCTTCAAACCTTGACTTCAGACGCGTCTGTAGTTCCCGTCACTTTTGAAGA 2503
Qy 1140 GAAGCGCATGGAAGAGGTTTTCGAGGATGACTGCTGCTGCAATTGACGTTGAGCGAAAA 1199
Db 2502 TATAGTGCCTTCTGTATACATAACCTTCGCGCATGGCACTCTTGAAAAAGTTCATGCCGTTT 2443
Qy 1200 CGCAGTTTACCATGATGATTCGGGAAGGTGTGGCCATGACAGCCCTTTTAAACGTTGAACTG 1259
Db 2442 CATATGATCCGGATAACGAGAAAGCAITTGAAACACCATGAGAGAAAGTAGTGACAAAGTGT 2383
Qy 1260 TTGTTTCAGCCACCTGGGATACCACTTCGTGCGGCTTTTTCGGACACAGTTTCGGATG 1319
Db 2382 TGGCCATGGAACAGGTAGTTTCCAGTAGTGCAATAAATTTAAGTTGTAAGCTTTCCGTA 2323
Qy 1320 GTCAGCCGAAGCGCATCAGCAACCCGAAACATACCGGCGACAGCGGAACCTGCCGTGCC 1379
Db 2322 TGTAGCATCACCTTCAACCTCTCCACTGACAGAAATTTGTGCCATTAACATCACCATC 2263
Qy 1380 GGTGTGAGATTAATGACAGCGGTGCGCGCTGGGATATTAAGTACAGAGAGCGGGTAT 1439
Db 2262 TAAATCAACAAGAAAT-----TGGGACAACTCCAGTGAAGTTTC 2224
Qy 1440 CTTGGCTGATCCGAGAAATGGAATGATACCCCGTGGTACCCGCGGCGCT 1499
Db 2223 TTCTCTTACTCATCGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTAAATTTGCGCGCT 2164
Qy 1500 TGGCGTAATCATGCTATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTTCCAC 1559
Db 2163 TGGCGTAATCATGCTATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTTCCAC 2104
Qy 1560 ACAACATACGAGCGGAAGCATAAAGTTAAAGCTGGGTCCTTAATCAGTGAGCTAAC 1619
Db 2103 ACAACATACGAGCGGAAGCATAAAGTTAAAGCTGGGTCCTTAATCAGTGAGCTAAC 2044
Qy 1620 TCACATTAATTTGGTTGGCTCACTGCGGCTTCCAGTCCGGAACCTGTGTCGACG 1679
Db 2043 TCACATTAATTTGGTTGGCTCACTGCGGCTTCCAGTCCGGAACCTGTGTCGACG 1984
Qy 1680 TGCATTAATGAATTCGGCAACCGCGGGAGAGCGGTTTGGGTATTTGGGCGCTCTTCCG 1739
Db 1983 TGCATTAATGAATTCGGCAACCGCGGGAGAGCGGTTTGGGTATTTGGGCGCTCTTCCG 1924
Qy 1740 CTTCTCGCTCACTGACTCGCTGCGTTCGGTTCGGTTCGGGCGAGCGGTATCAGCTC 1799
Db 1923 CTTCTCGCTCACTGACTCGCTGCGTTCGGTTCGGTTCGGGCGAGCGGTATCAGCTC 1864
Qy 1800 ACTCAAGCGGTTATACGGTTATCCAGAGATCAGGGATTAACGAGGAAGACATGT 1859
Db 1863 ACTCAAGCGGTTATACGGTTATCCAGAGATCAGGGATTAACGAGGAAGACATGT 1804
Qy 1860 GAGCAAAAGCGCAGCAAAAGCGCAGGAAACCGTAAAGAGCGCGTTTGGTGGGTTTTTCC 1919
Db 1803 GAGCAAAAGCGCAGCAAAAGCGCAGGAAACCGTAAAGAGCGCGTTTGGTGGGTTTTTCC 1744
Qy 1920 ATAGGCTCCGCCCTCTGACGAGCATCACAAAAATCAGCTCAAGTCAAGGTGGCGAA 1979
Db 1743 ATAGGCTCCGCCCTCTGACGAGCATCACAAAAATCAGCTCAAGTCAAGGTGGCGAA 1684
Qy 1980 ACCGACAGGACTATAAAGATACAGCGGTTTCCCGCTGGAGAGCTCCCTCGTGGCTCTC 2039
|||||

Db 1683 ACCGACAGGACTATAAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTC 1624
Qy 2040 CTGTTCCGACCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTCGGGAAGCGTGG 2099
Db 1623 CTGTTCCGACCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTCGGGAAGCGTGG 1564
Qy 2100 CGCTTCTCATAGCTCAGCTGTAGTATCTCAGTTTCGGTGTAGGTTCGCTCCCAAGC 2159
Db 1563 CGCTTCTCATAGCTCAGCTGTAGTATCTCAGTTTCGGTGTAGGTTCGCTCCCAAGC 1504
Qy 2160 TGGGCTGTGTGACGAAACCCCGTTTACGCCGACCGCTGCGCTTTATCCGGTAACATTC 2219
Db 1503 TGGGCTGTGTGACGAAACCCCGTTTACGCCGACCGCTGCGCTTTATCCGGTAACATTC 1444
Qy 2220 GTCTTGAGTCCAAACCGGTAAGACAGACTTATCGCACTGGCAGCAGCACTGGTAACA 2279
Db 1443 GTCTTGAGTCCAAACCGGTAAGACAGACTTATCGCACTGGCAGCAGCACTGGTAACA 1384
Qy 2280 GGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT 2339
Db 1383 GGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT 1324
Qy 2340 ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTACCTTCG 2399
Db 1323 ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTACCTTCG 1264
Qy 2400 GAAAAAGAGTTGGTACTCTTGATCCGCAAAACAAACCAACCGCTGCTGAGCGTGGTTTTT 2459
Db 1263 GAAAAAGAGTTGGTACTCTTGATCCGCAAAACAAACCAACCGCTGCTGAGCGTGGTTTTT 1204
Qy 2460 TTGTTTGAAGCAGCAGGATTACGCGCAGAAAAAAGGATCTCAAGAGATCCTTTGATCT 2519
Db 1203 TTGTTTGAAGCAGCAGGATTACGCGCAGAAAAAAGGATCTCAAGAGATCCTTTGATCT 1144
Qy 2520 TTCTTACGGGCTGACGCTCAGTGGAACGAAAACTCAGTTAAAGGATTTTGGTCATGA 2579
Db 1143 TTCTTACGGGCTGACGCTCAGTGGAACGAAAACTCAGTTAAAGGATTTTGGTCATGA 1084
Qy 2580 GATTATCAAAAAAGGATCTTACCTAGATCCTTTTAAATTTAAATTTAAATTTAAATCAA 2639
Db 1083 GATTATCAAAAAAGGATCTTACCTAGATCCTTTTAAATTTAAATTTAAATTTAAATCAA 1024
Qy 2640 TCTAAAGTATATATAGGTAACCTTGGTCTGACAGATTACCAATGCTTTAATCAGTGAGCAC 2699
Db 1023 TCTAAAGTATATATAGGTAACCTTGGTCTGACAGATTACCAATGCTTTAATCAGTGAGCAC 964
Qy 2700 CTATCTCAGCGATCTGTCTATTTTCTCATCATAGTTCGCTGACTCCCGCTCGTGTAGA 2759
Db 963 CTATCTCAGCGATCTGTCTATTTTCTCATCATAGTTCGCTGACTCCCGCTCGTGTAGA 904
Qy 2760 TAACTACGATACGGAGGCTTACCATCTGGCCCCCAGTCTGCAATGATACCGCGAGACC 2819
Db 903 TAACTACGATACGGAGGCTTACCATCTGGCCCCCAGTCTGCAATGATACCGCGAGACC 844
Qy 2820 CACGCTCACCGGCTCCAGATTTTATCAGCAATAAACACAGCAGCCGGAAGGCGGAGCGCA 2879
Db 843 CACGCTCACCGGCTCCAGATTTTATCAGCAATAAACACAGCAGCCGGAAGGCGGAGCGCA 784
Qy 2880 GAAGTGTCTTCAACTTTTATCCGCTCCATCCAGTCTATTAATTTCTGCGGGAAGCTA 2939
Db 783 GAAGTGTCTTCAACTTTTATCCGCTCCATCCAGTCTATTAATTTCTGCGGGAAGCTA 724
Qy 2940 GAGTAAGTAGTTTCGCCAGTTAATAGTTTTCGCAACCGTTTTCGCAATTTGCTACAGGATCG 2999
Db 723 GAGTAAGTAGTTTCGCCAGTTAATAGTTTTCGCAACCGTTTTCGCAATTTGCTACAGGATCG 664
Qy 3000 TGGTGTACGCTCGCTGTTTGGTATGGCTTCAATTAAGTCCGCTTCCCAACGATCAAGGC 3059
Db 663 TGGTGTACGCTCGCTGTTTGGTATGGCTTCAATTAAGTCCGCTTCCCAACGATCAAGGC 604
Qy 3060 GAGTTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTTCTCGGCTCCCGATCG 3119
Db 603 GAGTTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTTCTCGGCTCCCGATCG 544
|||||

QY 3120 TTGTGAGAGTAAGTTGGCGCAGTGTATCACTCATGCTTATGGCAGCACTGCATAATT 3179
DB 543 TTGTGAGAGTAAGTTGGCGCAGTGTATCACTCATGCTTATGGCAGCACTGCATAATT 484
QY 3180 CTCTACTGTCATGCGCATCCGTAAGATGCTTTCTGTGACTGGTGAGTACTCAACCAAGT 3239
DB 483 CTCTACTGTCATGCGCATCCGTAAGATGCTTTCTGTGACTGGTGAGTACTCAACCAAGT 424
QY 3240 CATTCTGAGAAATAGTATGCGCGCAGCAGTGTCTTTGCCGCGGTCAATACGGGATA 3299
DB 423 CATTCTGAGAAATAGTATGCGCGCAGCAGTGTCTTTGCCGCGGTCAATACGGGATA 364
QY 3300 ATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAAACGTTCTTCGGGGC 3359
DB 363 ATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAAACGTTCTTCGGGGC 304
QY 3360 GAAACTCTCAAGGATCTTACCCTGTTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 3419
DB 303 GAAACTCTCAAGGATCTTACCCTGTTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 244
QY 3420 CCAACTGATCTTCAGCATCTTTTACTTTTCCAGCGTTTCTGGTGAGCAAAAACAGGAA 3479
DB 243 CCAACTGATCTTCAGCATCTTTTACTTTTCCAGCGTTTCTGGTGAGCAAAAACAGGAA 184
QY 3480 GGCAGAAATGCCGCAAAAAGGAATAAGGGCGACACGGAATGTTGAATACTCATCTCT 3539
DB 183 GGCAGAAATGCCGCAAAAAGGAATAAGGGCGACACGGAATGTTGAATACTCATCTCT 124
QY 3540 TCCCTTTTCAATATTAATGAAGCATTTATCAGGGTTTATGCTCATGAGCGGATACATAT 3599
DB 123 TCCCTTTTCAATATTAATGAAGCATTTATCAGGGTTTATGCTCATGAGCGGATACATAT 64
QY 3600 TTGAATGATTTAGAAAAATAAACAATAGGGTTCCCGGCACATTTCCCGGAAAAGTGC 3659
DB 63 TTGAATGATTTAGAAAAATAAACAATAGGGTTCCCGGCACATTTCCCGGAAAAGTGC 4
QY 3660 CAC 3662
DB 3 CAC 1

RESULT 5
ADSL17192/c
ID ADSL17192 standard; DNA; 3637 BP.
XX
AC ADSL17192;
XX
DT 02-DEC-2004 (first entry)
XX
DE pBSC3BFP plasmid DNA encoding A. victoria BFP Cycle 3 gene.
XX
KW Polymerase; BFP; jellyfish; blue fluorescent protein; chimeric; gene; ds.
XX
OS Aequorea victoria.
OS Chimeric.
OS Unidentified.
XX US2004180352-A1.
XX
PD 16-SEP-2004.
XX
PF 08-AUG-2003; 2003US-00637758.
XX
PR 08-AUG-2002; 2002US-0402342P.
XX (LARG-) LARGE SCALE BIOLOGY CORP.
XX Padgett HS, Lindbo JA, Fitzmaurice WP, Vaewhongs AA;
XX WPI; 2004-667658/65.
XX
PT In vitro method of making sequence variants from heteroduplex

PT polynucleotide, involves combining polynucleotide with agent having
PT polymerase and strand cleavage activities for sufficient time for
PT percentage of complementarity to increase.
XX Example 18; SEQ ID NO 32; 81pp; English.
PS The invention relates to an in vitro method of making sequence variants
CC from heteroduplex polynucleotide which involves combining polynucleotide
CC with an agent having polymerase and strand cleavage activities for
CC sufficient time for the percentage of complementarity to be increased
CC within the heteroduplex. The method is useful for performing in vitro
CC method of making sequence variants from one or more heteroduplex
CC polynucleotide. It is useful in increasing diversity in a population of
CC sequences and obtaining a polynucleotide encoding a desired functional
CC property. The present sequence is a plasmid DNA encoding Aequorea
CC victoria blue fluorescent protein (BFP) Cycle 3 gene. This sequence is
CC used to illustrate the method of the invention.
XX
SQ Sequence 3637 BP; 942 A; 891 C; 870 G; 934 T; 0 U; 0 Other;
Query Match 72.8%; Score 2665.4; DB 13; Length 3637;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;
QY 1 CTAAATTGTAAGCGTTAATAATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60
DB 3637 CTAAATTGTAAGCGTTAATAATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 3578
QY 61 ATTTTAAACNATAGGCGGAAATCGGCAAAATCCCTTATAAATCAAAAGATAGACCGA 120
DB 3577 ATTTTAAACNATAGGCGGAAATCGGCAAAATCCCTTATAAATCAAAAGATAGACCGA 3518
QY 121 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAAGTGGACTC 180
DB 3517 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAAGTGGACTC 3458
QY 181 CAACTGTCAAAGGGCGAAAACCGCTCTATCAGGGCGATGGCCACTACGTGAAACCATCACC 240
DB 3457 CAACTGTCAAAGGGCGAAAACCGCTCTATCAGGGCGATGGCCACTACGTGAAACCATCACC 3398
QY 241 CTAAATCAAGTTTGTGGGTGAGTGCCTTAAGACACTAAATCGGAACCTTAAAGGGAG 300
DB 3397 CTAAATCAAGTTTGTGGGTGAGTGCCTTAAGACACTAAATCGGAACCTTAAAGGGAG 3338
QY 301 CCCCCGATTTAGAGCTTCACGCGGAAAGCGCGCAAACTGGCGAGAAAGAGGGAAGAA 360
DB 3337 CCCCCGATTTAGAGCTTCACGCGGAAAGCGCGCAAACTGGCGAGAAAGAGGGAAGAA 3278
QY 361 AGCGAAAGAGCGGCGCTAGGCGCTGGCAAGTGTAGCGGTACGCTGCGGTAAACCAAC 420
DB 3277 AGCGAAAGAGCGGCGCTAGGCGCTGGCAAGTGTAGCGGTACGCTGCGGTAAACCAAC 3218
QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCCATTCGCCATTGAGCTGCG 480
DB 3217 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCCATTCGCCATTGAGCTGCG 3158
QY 481 CAACTGTTGGGAAGGCGCATCGGTGCGGCGCTCTTTCGCTATTACGCCAGTGGCGAAAGG 540
DB 3157 CAACTGTTGGGAAGGCGCATCGGTGCGGCGCTCTTTCGCTATTACGCCAGTGGCGAAAGG 3098
QY 541 GGGATGTGCTGCAAGGCGGATTAAGTTGGGTAAACGCCAGGGTTTTCCAGTACGACGTTG 600
DB 3097 GGGATGTGCTGCAAGGCGGATTAAGTTGGGTAAACGCCAGGGTTTTCCAGTACGACGTTG 3038
QY 601 TAAACGACGCGCAGTGAGCGGCGCTCGTTCACTTCACTGTTTTTGAACCCGCTGAGGACGG 660
DB 3037 TAAACGACGCGCAGTGAGCGGCGCTCGTTCACTTCACTGTTTTTGAACCCGCTGAGGACGG 2978
QY 661 GCAGACTCGCGTGCAAAATGTGTTTTACAGCGTGTATGAGCAGATGAAGATGCTCGACAC 720
DB 2977 CGCGGTGCGGCGCTCTAGAACTAGTGGATCCCGCGGCTGCGAGGAATCTTATTGTT 2918
QY 721 GCTGCAGAACACGCGAGCTAGATTAAACCTAGAAAAGATAATCATATTGTGACGTGTTAA 780

Db 2917 ATAGTTCAATCCATGCCATGTGTAATCCACGACGAGTTACAACTCAAGAGACCATGT 2858
QY 781 AGATAATC-ATGCGTAAATTTAGCATGGATCTGTAAATACGACTCACTATAGGCGGAA 839
Db 2857 GGTACGCTTTTGGTGGATCTTTGAAAGGCGAGATTGTGTCGACAGTAATGTTGT 2798
QY 840 TTGGGTACCGGGCCCCCTCGAGGTGCGAGGTATCGATAAGCTTGATATCGAATTCCTG 899
Db 2797 CTGTGTAAGGACAGGGCCATCCCAATTGGAGTATTTTGTGATAATGGTCTGTAGTT 2738
QY 900 CAGCCCGGGGATCCACTAGTTCTAGAGCGCGCCACCGCGTGGAGTCCAGCTTTTG 959
Db 2737 GAACGATCCATCTTCAATGTTGTGGGAATTTTGAAGTTAGTTTGAATCCATTCCTTT 2678
QY 960 TTCCCTTTAGTGAAGGTTAATTAGATCCCATGCGTCAATTTTACGAGACTATCTTTCTA 1019
Db 2677 GTTTGTCTGCGGTGATATACATTGTGTAGTTAAAGTTGTACTCGAGTTGTGTCGA 2618
QY 1020 GGGTTAATCTAGCTGATCAGGATCATATCGTGGGTCTTTTTCGGGTCTAGTCAATCGC 1079
Db 2617 GAATGTTTCCATCTTCTTTAAATCAATACCTTTTAACTCGATAC-----GATTAAAG 2563
QY 1080 CCAAGCTGGCGTATCTGGGCATCGGGAGGAAGCCGCTTTTCCCGGAGGTT 1139
Db 2562 GGTATCACCTTCAACTTGACTTCAGACGCTTGTAGTTCCCGTCACTTTTGAAGA 2503
QY 1140 GAAGCGCATGAAAGAGTTTGCAGGATGACTGCTGCTGCAATTGACCTTGAGCGAAAA 1199
Db 2502 TATAGTGGTCTCTGTATACAACTTTCGGGATGGCACTTTGAAAGTCAATCGCTT 2443
QY 1200 CGACGTTTACATGATGATTCGGGAAGGTGGCCATGACGCCCTTTTAAACGTTGAAGT 1259
Db 2442 CATATGATCCGATACGAGAAAGCAITGCAACCATCAGAGAAAGTAGTGACAAAGT 2383
QY 1260 TTGTTTACGCCACCTGGGATACCACTGCTGCGGCTTTTCCGGACAGTTTCCGGATG 1319
Db 2382 TGGCCATGGAACAGGTAGTTTCCAGTAGTGCAATAAATTTAAGTGTAAAGCTTTCCGTA 2323
QY 1320 GTACGCCGAAGCGCATCAGCAACCCGAAACATACCGGGACAGCGGAACCTGCCGTGCC 1379
Db 2322 TGTAGCATCACCTTCACTCCCTCCACTGACAGAAATTTGTGCCATTAAATCACCATC 2263
QY 1380 GGTGTGCAATTAATGACAGCGGTGGCGCTGGGATATTAAGTACAGAGAGCGGTAT 1439
Db 2262 TAATTCACAAAGAAAT-----TGGACAACTCCAGTGAAAGATTC 2224
QY 1440 CCTGCTGGATGCGCAGAAATGACATGGATACCCGCTGAGTTACCCGCGGCGCGCT 1499
Db 2223 TTTCTCTTTACTCATCGGTACCCAGCTTTTGTCTTGTAGGAGGTTAAATGCGCGCT 2164
QY 1500 TGGCGTAAATCATGCTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTCAC 1559
Db 2163 TGGCGTAAATCATGCTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTCAC 2104
QY 1560 ACAACATAGAGCCGGAAGATTAAGCTTGAAGCTTGGGGTGCTTAATGAGTGAGCTAAC 1619
Db 2103 ACAACATAGAGCCGGAAGATTAAGCTTGAAGCTTGGGGTGCTTAATGAGTGAGCTAAC 2044
QY 1620 TCACATTAATGCTTGGCTCACTGCGCTTTCAGTCCGGAACCTGCTGTCGCGAG 1679
Db 2043 TCACATTAATGCTTGGCTCACTGCGCTTTCAGTCCGGAACCTGCTGTCGCGAG 1984
QY 1680 TGCATTAATGAATTCGGCACAACGCGGGGAGAGCGGTTTGGGTATTTGGCGCTCTTTCCG 1739
Db 1983 TGCATTAATGAATTCGGCACAACGCGGGGAGAGCGGTTTGGGTATTTGGCGCTCTTTCCG 1924
QY 1740 CTTTCTCGCTCACTGACTCGTGGCTCGGTTCGTTGCGGCGAGCGGTATCAGCTC 1799
Db 1923 CTTTCTCGCTCACTGACTCGTGGCTCGGTTCGTTGCGGCGAGCGGTATCAGCTC 1864
QY 1800 ACTCAAGCGGTAAATACGGTTATCACAAGATCAGGGATTAACGAGAAAGACATGT 1859
|||

Db 1863 ACTCAAGCGGGTAATACGGTTATCCACAGAATCAGGGATTAACGAGGAAGAACATGT 1804
QY 1860 GAGCAAAAAGCCAGCAAAAAGCCAGAACCGTAAAAAAGCCGCTTGTGCGGTTTTTCC 1919
Db 1803 GAGCAAAAAGCCAGCAAAAAGCCAGAACCGTAAAAAAGCCGCTTGTGCGGTTTTTCC 1744
QY 1920 ATAGCTCCGCCCCCTGACAGGATCACAAAAATCGACGCTCAAGTCAAGAGTGGCGAA 1979
Db 1743 ATAGGCTCCGCCCCCTGACAGGATCACAAAAATCGACGCTCAAGTCAAGAGTGGCGAA 1684
QY 1980 ACCGACAGGACTATAAAGATACCAAGCGTTTTCCCTCGAAGCTCCCTCGTGGCTCTC 2039
Db 1683 ACCGACAGGACTATAAAGATACCAAGCGTTTTCCCTCGAAGCTCCCTCGTGGCTCTC 1624
QY 2040 CTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTTCTCTCCGGAAGCGTGG 2099
Db 1623 CTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTTCTCTCCGGAAGCGTGG 1564
QY 2100 CGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGCTGAGGTGCTTCCGCTCAAGC 2159
Db 1563 CGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGCTGAGGTGCTTCCGCTCAAGC 1504
QY 2160 TGGGCTGTGTGCACGAAACCCCGTTTCAGCCCGACCGCTTCGCTTATCCGGTAACTATC 2219
Db 1503 TGGGCTGTGTGCACGAAACCCCGTTTCAGCCCGACCGCTTCGCTTATCCGGTAACTATC 1444
QY 2220 GTCTTGAFTCCAAACCGGTAAGACAGCTTATCGCACTGGCAGCAGCACTGGTAACA 2279
Db 1443 GTCTTGAFTCCAAACCGGTAAGACAGCTTATCGCACTGGCAGCAGCACTGGTAACA 1384
QY 2280 GATTTAGCAGAGGAGGTATGTAGGGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAAC 2339
Db 1383 GATTTAGCAGAGGAGGTATGTAGGGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAAC 1324
QY 2340 ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACCTTCG 2399
Db 1323 ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACCTTCG 1264
QY 2400 GAAAAAGAGTTGGTAGCTCTTGATCCGCAAAACAAACCCGCTGGTAGCGGTGTTTTT 2459
Db 1263 GAAAAAGAGTTGGTAGCTCTTGATCCGCAAAACAAACCCGCTGGTAGCGGTGTTTTT 1204
QY 2460 TTGTTTGAAGCAGCAGATTTAGCCGAGAAAGGATCTCAAGAGATCTCTTGTATCT 2519
Db 1203 TTGTTTGAAGCAGCAGATTTAGCCGAGAAAGGATCTCAAGAGATCTCTTGTATCT 1144
QY 2520 TTTCTACGGGCTCTGACGCTCAGTGGAAACGAAACCTCAGCTTAAAGGATTTTGGTCA 2579
Db 1143 TTTCTACGGGCTCTGACGCTCAGTGGAAACGAAACCTCAGCTTAAAGGATTTTGGTCA 1084
QY 2580 GATTTCAAAAAGGATCTTACCTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAA 2639
Db 1083 GATTTCAAAAAGGATCTTACCTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAA 1024
QY 2640 TCTAAAGTATATATGATTAAGTAACTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGG 2699
Db 1023 TCTAAAGTATATATGATTAAGTAACTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGG 964
QY 2700 CTATCTCAGGATCTGTCTATTTTCTATCATCATAGTTGCTGCTGCTGCTGCTGCTGCT 2759
Db 963 CTATCTCAGGATCTGTCTATTTTCTATCATCATAGTTGCTGCTGCTGCTGCTGCTGCT 904
QY 2760 TAACTACGATCGGGAGGCTTACCATCTGGCCCCCAGTCTGCAATGATACCGCGAGACC 2819
Db 903 TAACTACGATCGGGAGGCTTACCATCTGGCCCCCAGTCTGCAATGATACCGCGAGACC 844
QY 2820 CAGCTCACCGGCTCCAGATTTTATCAGCAATAAACCCAGCAGCCGGAAGGCGGAGCGCA 2879
Db 843 CAGCTCACCGGCTCCAGATTTTATCAGCAATAAACCCAGCAGCCGGAAGGCGGAGCGCA 784
QY 2880 GAAGTGGTCTCTCAACTTTTATCCGCTCCATCCAGTCTAATTTGTTCCCGGGAAGCTA 2939
Db 783 GAAGTGGTCTCTCAACTTTTATCCGCTCCATCCAGTCTAATTTGTTCCCGGGAAGCTA 724
|||

QY 2940 GAGTAAGTAGTTTCGCCAGTTAAATAGTTTTCGCCAAAGTGTGTTGCCATTGCTACAGGCATCG 2999
DB 723 GAGTAAGTAGTTTCGCCAGTTAAATAGTTTTCGCCAAAGTGTGTTGCCATTGCTACAGGCATCG 664
QY 3000 TGGTGTACAGCTCGTGGTGGTATGGCTTCAATTCAGCTCCGGTTCCTCCCAACGATCAAGGC 3059
DB 663 TGGTGTACAGCTCGTGGTGGTATGGCTTCAATTCAGCTCCGGTTCCTCCCAACGATCAAGGC 604
QY 3060 GAGTTACATGATCCGCCATGTTGTGCAAAAAGCGGTAGCTCTCTCGGTCTCCGATCG 3119
DB 603 GAGTTACATGATCCGCCATGTTGTGCAAAAAGCGGTAGCTCTCTCGGTCTCCGATCG 544
QY 3120 TTGTGAGAAGTAAGTTGGCCGAGTGTATCACTCATGCTTATGCGACGACATGCATAATT 3179
DB 543 TTGTGAGAAGTAAGTTGGCCGAGTGTATCACTCATGCTTATGCGACGACATGCATAATT 484
QY 3180 CTCCTACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGATCGTGGTACTCAACCAAGT 3239
DB 483 CTCCTACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGATCGTGGTACTCAACCAAGT 424
QY 3240 CATTCGAGANTAGTGTATCGCGCACCGAGTTGCTTTCGCCGGGTCATACGGGATA 3299
DB 423 CATTCGAGANTAGTGTATCGCGCACCGAGTTGCTTTCGCCGGGTCATACGGGATA 364
QY 3300 ATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCAATTGGAAGCGTCTTCGCGGC 3359
DB 363 ATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCAATTGGAAGCGTCTTCGCGGC 304
QY 3360 GAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGAC 3419
DB 303 GAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGAC 244
QY 3420 CCAACTGATCTTCAGCATCTTTTACTTTTCACAGCGTTTCTGGGTGAGCAAAAACAGAA 3479
DB 243 CCAACTGATCTTCAGCATCTTTTACTTTTCACAGCGTTTCTGGGTGAGCAAAAACAGAA 184
QY 3480 GGCAAAATGCCCAAAAAGGGAATAAGGGGCACACGGAATGTTGAATCTCATCTCT 3539
DB 193 GGCAAAATGCCCAAAAAGGGAATAAGGGGCACACGGAATGTTGAATCTCATCTCT 124
QY 3540 TCCTTTTCAATATTATTGAAGCATTTATCAGGGTTATGTTCTCATGAGCGGATACATAT 3599
DB 123 TCCTTTTCAATATTATTGAAGCATTTATCAGGGTTATGTTCTCATGAGCGGATACATAT 64
QY 3600 TTGAATGATTTAGAAAAATAAACAATAGGGGTTCCGGGCACATTTCCCGGAAAAGTGC 3659
DB 63 TTGAATGATTTAGAAAAATAAACAATAGGGGTTCCGGGCACATTTCCCGGAAAAGTGC 4
QY 3660 CAC 3662
DB 3 CAC 1
RESULT 6
ID ABX14571/c
XX ABX14571 standard; DNA; 3637 BP.
AC ABX14571;
DT 14-MAR-2003 (first entry)
DE Plasmid pBSC3GFP DNA.
XX Circular; cyclic; heteroduplex; T4 DNA polymerase; T4 DNA ligase; CEL I;
KW population diversity; exonuclease; complementarity; ds.
XX Synthetic.
XX OS
XX WO200279468-A2.
XX PN
XX PD
XX 10-OCT-2002.

PF 01-FEB-2002; 2002WO-US003055.
XX
PR 02-FEB-2001; 2001US-0266386P.
PR 14-FEB-2001; 2001US-0268785P.
XX
PA (LARG-) LARGE SCALE BIOLOGY CORP.
XX
PI Padgett HS, Fitzmaurice WP, Lindo JA;
XX WPI; 2003-092898/08.
DR
XX
XX Increasing complementarity by making sequence variants from heteroduplex
PT polynucleotides, useful for producing improved gene products from
PT randomly mutated genes.
PT
PS Example 5; Page 99-100; 105pp; English.
XX
CC The invention relates to an in vitro method of making sequence variants
CC from at least one heteroduplex polynucleotide that has at least two-non
CC complementary nucleotide base pairs. The method comprises combining a
CC heteroduplex polynucleotide with CEL I, T4 DNA polymerase and T4 DNA
CC ligase, or an agent or agents with exonuclease activity and allowing
CC sufficient time for the percentage of complementarity to increase, where
CC one or more variants are made. The method can also be used to increase
CC diversity in a population of sequences. The method is useful for
CC producing improved gene products from randomly mutated genes or from
CC expression improved gene products from animal, fungal, yeast or bacterial
CC expression vector, and provides a high-efficiency recovery of
CC recombinants for generating novel polynucleotides from parental templates
CC that are more diverse and with a lower percentage of sequence identity.
CC This sequence represents plasmid DNA, used in heteroduplex substrate
CC preparation
XX
SQ Sequence 3637 BP; 942 A; 892 C; 870 G; 933 T; 0 U; 0 Other;
Query Match 72.8%; Score 2664.8; DB 8; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;
QY 1 CTAATTTCTAGCGTTAATATTTTCTTAAATTCGCGTTTAAATTTTCTTAAATCAGCTC 60
DB 3637 CTAATTTCTAGCGTTAATATTTTCTTAAATTCGCGTTTAAATTTTCTTAAATCAGCTC 3578
QY 61 ATTTTAAACCAATAGGCGGAAATCGCAAAATCCCTTATAAATCAAAAGATAGACCGA 120
DB 3577 ATTTTAAACCAATAGGCGGAAATCGCAAAATCCCTTATAAATCAAAAGATAGACCGA 3518
QY 121 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGTCCACTATTATAAGACGTGGACTC 180
DB 3517 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGTCCACTATTATAAGACGTGGACTC 3458
QY 181 CAACGTCAAGGGCGAATAACCGTCTATCAGGGCGATGGCCACTACGTGAACCATCACC 240
DB 3457 CAACGTCAAGGGCGAATAACCGTCTATCAGGGCGATGGCCACTACGTGAACCATCACC 3398
QY 241 CTAATCAAGTTTCTTGGGGTTCGAGTTCGCGTAAAGACACTAAATTCGGAACCTTAAAGGGAG 300
DB 3397 CTAATCAAGTTTCTTGGGGTTCGAGTTCGCGTAAAGACACTAAATTCGGAACCTTAAAGGGAG 3338
QY 301 CCCCCGATTAGAGCTTGACGGGGAAAGCCGCGCAACGTGGCGAGAAAGGAAGGAAGAA 360
DB 3337 CCCCCGATTAGAGCTTGACGGGGAAAGCCGCGCAACGTGGCGAGAAAGGAAGGAAGAA 3278
QY 361 AGCGAAGAGCGGCGCTAGGCGCTGCGAAGTGTAGCGGTACGCTGCGGTAAACAC 420
DB 3277 AGCGAAGAGCGGCGCTAGGCGCTGCGAAGTGTAGCGGTACGCTGCGGTAAACAC 3218
QY 421 CACACCCCGCGCTTAAATGCGCCCTACAGGGCGGTCCTCCATTCGCCATTTCAGGCTGCG 480
DB 3217 CACACCCCGCGCTTAAATGCGCCCTACAGGGCGGTCCTCCATTCGCCATTTCAGGCTGCG 3158
QY 481 CAACTGTTGGGAAGGGCGATCGGTGCGGCGCTCTTTCGCTATTACGCCAGCTGCGGAAAGG 540
|||||

Db 3157 CAACTGTTGGGAAGGCGATCGGTGGGGCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098
QY 541 GGGATGTGTGCAAGCGGATTAAGTTGGGTAAAGCGCAGGGTTTCCAGTCACGAGCTTG 600
Db 3097 GGGATGTGTGCAAGCGGATTAAGTTGGGTAAAGCGCAGGGTTTCCAGTCACGAGCTTG 3038
QY 601 TAAAGCAGCGGCGAGTGGGGCTCGTTTCATTACGTTTTTGAACCCGTGGAGACGG 660
Db 3037 TAAAGCAGCGGCGAGTGGGGCGGTAATACGACTCACTATAGGGCGAATTTGGAGCTCCA 2978
QY 661 GCAGACTCGCGGTGCAAAATGTGTTTACAGCGTGAATGGAGCAGATGAAGATGCTCGACAC 720
Db 2977 CCGCGTGGCGG-----CGCTCTAGAACTAGTGATCCCC--CGG 2939
QY 721 GTGCGAGACACGCGACTAGATTAAACCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
Db 2938 GTGCGAGGAATCTTATTGTTGATAGTTCAATCCATGCCATGTGTAATCCCGCAGCAGTTA 2879
QY 781 AGATAATCATGCTAAATTTGACGATGGATCTGTAATACGACTCACTATAGGGCGAAT 840
Db 2878 CAAACTCAAGAGGACCATGTGGTCACGCTTTTCGTTGGGATCTTTCGAAAGGGCAGATT 2819
QY 841 TGGGTACCGGGCCCCCTCGAGTCCGAGTCGACGGTATCGATAAGCTTGATATCGAATTCCTGC 900
Db 2818 GTGTCGACAGGTAAATGGTTGTCTGGTTAAAGACAGGGCCATCGCCAAATTGAGTATTT 2759
QY 901 AGCCCGGGGATCCACTAGTTCTAGAGCGCGCGCCACCGCGGTGGAGTCCAGCTTTTGT 960
Db 2758 GTTGATAATGGTCTGTAGTTGAACGGATCCATCTTCAATGTTGTGGCGAATTTTGAAGT 2699
QY 961 TCCCTTTAGTAGGGTTAATAGATCCCATGGTCAATTTTACCGCAGACTATCTTCTAG 1020
Db 2698 TAGCTTTGATCCATTCCTTTGTTGTCCTGCGGTGATGATACATTTGTGTGAG-TTATAG 2640
QY 1021 GGTAACTAGCTGCATCAGGATCATCTCGTGGTCTTTTTTCGGCTCAGTCATCGCC 1080
Db 2639 TTGTACTCAGTTTGTGTCGGAATGTTTTCATCTCTTTAAATCAATACCTTTTAAAC 2580
QY 1081 CAACTGGCGCTATCTGGGCATCGGGAGGAGAAAGCCGTGCTTTTCCCGCAGGTTG 1140
Db 2579 TCGATACGATTAAACAGGGTATCACCTTCAAACTTGACTTCAGCAGCGCTTGTAGTTT 2520
QY 1141 AAGCGCATGGAAGGTTTGGCGAGGATGACTGTGCTGCATGACGTTGAGCGGAAC 1200
Db 2519 CGGTCACTTTGAAAGATATAGTGTCTTCTGTACATAACCTTCGGGCATGGCACTCTTG 2460
QY 1201 GCACGTTTACCATGATTTCCGGAGGTGTGCGCATGACCGCTTTTACGGTGAAGTCT 1260
Db 2459 AAAAGTCAATCGCGTTTCATATGATCCGGATTAACGGGAAAGCATTTGAACACCAATAGAG 2400
QY 1261 TCGTTTCAGGCCACCTGGGATACCAAGTTCTGCGGCTTTTTCGGGACACAGTTCCGGATGG 1320
Db 2399 AAGTAGTGACAAGTGTGGCCATGGAACAGGTAGTTTCCAGTAGTGCAATAAATTTA 2340
QY 1321 TCAGCCGGAAGCGCATCAGAACCCGAAACAATACCGGCGACAGCGGAACTCGCGTGGCG 1380
Db 2339 AGGGTAAGCTTTCCGATATAGCATCACCTTCACTCTCCATCGACGAAATTTGTGC 2280
QY 1381 GTGTGCAAGTAAATGACAGCGGTGGCGCTGGGATTTACGTGACGAGGACGGGTATC 1440
Db 2279 CCAITTAACATCACTAATTTCAACAAGAAATTTGGGACAACTCCAGTGAAGATTTCTTCT 2220
QY 1441 CTGGGTGGATGCGCGAAGATGACATGGATACCCCGTGAGTTTACCCGCGGGCGCGCTT 1500
Db 2219 C---CTTTACTCATCGTATACCGAGCTTTTGTTCCTTTTAGTGAGGGTTAATTGGCGCTT 2163
QY 1501 GCGTAATCATGGTCAATGCTGTTTCTGTGTGAATTTGATTCGCTCACAATTTCCACA 1560
Db 2162 GCGTAATCATGGTCAATGCTGTTTCTGTGTGAATTTGATTCGCTCACAATTTCCACA 2103
QY 1561 CAACTACGAGCGGAGCATAAAGGTAAAGCCTGGGTGCTTAATGATGAGTCACT 1620
Db 2102 CAACTACGAGCGGAGCATAAAGGTAAAGCCTGGGTGCTTAATGATGAGTCACT 2043

QY 1621 CACATTAATTGGTTGGCTCATCTGCCCGCTTTTCAGTCGGGAAACCTGTGTGTGCCAGCT 1680
Db 2042 CACATTAATTGGTTGGCTCATCTGCCCGCTTTTCAGTCGGGAAACCTGTGTGTGCCAGCT 1983
QY 1681 GCATTAATCAATCGCCAAACGCGCGGGAGAGCGGTTTTCGTTATTTGGGCGCTCTTCCGC 1740
Db 1982 GCATTAATCAATCGCCAAACGCGCGGGAGAGCGGTTTTCGTTATTTGGGCGCTCTTCCGC 1923
QY 1741 TTTCTCGCTCACTCACTCGCTCGCTCGCTCGTTTCGGCTCGCGCAGCGGTATCAGTCA 1800
Db 1922 TTTCTCGCTCACTCACTCGCTCGCTCGTTTCGGCTCGCGCAGCGGTATCAGTCA 1863
QY 1801 CTCAAAGGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGCGAGGAAAGAAATGTG 1860
Db 1862 CTCAAAGGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGCGAGGAAAGAAATGTG 1803
QY 1861 ASCAAAGGCCAGCAAAAGGCCAGGAACCGTAAAGAGCGGTTCTCGCGTTTTTCCA 1920
Db 1802 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAGAGCGGTTCTCGCGTTTTTCCA 1743
QY 1921 TAGGCTCCGCCCTCTGACGAGCATCAAAAATCCAGCTCAAGTCAGAGGTGGGAAA 1980
Db 1742 TAGGCTCCGCCCTCTGACGAGCATCAAAAATCCAGCTCAAGTCAGAGGTGGGAAA 1683
QY 1981 CCCGACAGGACTATAAAGATACAGGGTTTCCCTCTGGAAGCTCCCTCGTGGCTCTCC 2040
Db 1682 CCCGACAGGACTATAAAGATACAGGGTTTCCCTCTGGAAGCTCCCTCGTGGCTCTCC 1623
QY 2041 TGTTCGAGCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAGCGTGGC 2100
Db 1622 TGTTCGAGCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAGCGTGGC 1563
QY 2101 GCTTTCTCATAGCTCACTAGTATCTCAGTTTCGGTGTAGGTCTTCGCTCCAAGCT 2160
Db 1562 GCTTTCTCATAGCTCACTAGTATCTCAGTTTCGGTGTAGGTCTTCGCTCCAAGCT 1503
QY 2161 GGGCTGTGTGACGAAACCCCTGTTACGCCACCGCTCGGCTTATTCGGTAACTATCG 2220
Db 1502 GGGCTGTGTGACGAAACCCCTGTTACGCCACCGCTCGGCTTATTCGGTAACTATCG 1443
QY 2221 TCTTGAGTCCAAACCCGTTAAGACAGCATTTATGCCACTGGCAGCAGCCACTGTTAAACAG 2280
Db 1442 TCTTGAGTCCAAACCCGTTAAGACAGCATTTATGCCACTGGCAGCAGCCACTGTTAAACAG 1383
QY 2281 GATTAGCAGCAGGATATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTA 2340
Db 1382 GATTAGCAGCAGGATATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTA 1323
QY 2341 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTTCGG 2400
Db 1322 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTTCGG 1263
QY 2401 AAAAAGTTGTGTAGCTCTTGATCCGGGAAACAACCCAGCTGGTGTAGCGGTGGTTTTT 2460
Db 1262 AAAAAGTTGTGTAGCTCTTGATCCGGGAAACAACCCAGCTGGTGTAGCGGTGGTTTTT 1203
QY 2461 TGTTCAGCAGCAGATTTACCGCAGAAAAAAGGATCTCAAGAGATCTCTTGTATCTT 2520
Db 1202 TGTTCAGCAGCAGATTTACCGCAGAAAAAAGGATCTCAAGAGATCTCTTGTATCTT 1143
QY 2521 TTTCTACGGGTCTGACGCTCAGTGGAAAGAAACTCAAGTTAAGGATTTTGGTCAATGAG 2580
Db 1142 TTTCTACGGGTCTGACGCTCAGTGGAAAGAAACTCAAGTTAAGGATTTTGGTCAATGAG 1083
QY 2581 ATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCAAT 1023
QY 2641 CTAAAGTATATAGTAAACTTTGTTCTGACAGTTTACCAATCTTAATCAGTGAGGCACC 2700
Db 1022 CTAAAGTATATAGTAAACTTTGTTCTGACAGTTTACCAATCTTAATCAGTGAGGCACC 963

QY	2701	TATCTCAGCGATCTGTCTATTTTCTGTTCACTCCATAGTTCCTGACTCCCGTCTGTTAGAT	2760
Db	962	TATCTCAGCGATCTGTCTATTTTCTGTTCACTCCATAGTTCCTGACTCCCGTCTGTTAGAT	903
QY	2761	AACTAGGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCCGAGACCC	2820
Db	902	AACTAGGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCCGAGACCC	843
QY	2821	AGCTCACCGGCTCCAGATTTATCAGCAATAAACACAGCCAGCGGAGGGCCGAGCGAG	2880
Db	842	AGCTCACCGGCTCCAGATTTATCAGCAATAAACACAGCCAGCGGAGGGCCGAGCGAG	783
QY	2881	AAGTGGTCTCGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTGCGGGAAGTAG	2940
Db	782	AAGTGGTCTCGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTGCGGGAAGTAG	723
QY	2941	AGTAAGTAGTTCGCCAGTTAATAGTTTGGCGAACGTTGTGTCATTTGTACAGGCGATCGT	3000
Db	722	AGTAAGTAGTTCGCCAGTTAATAGTTTGGCGAACGTTGTGTCATTTGTACAGGCGATCGT	663
QY	3001	GGTGTACGCTCGTCTGTTGGTATGCTTCAATTCAGCTCCGGTTCCTCAACGATCAAGGCG	3060
Db	662	GGTGTACGCTCGTCTGTTGGTATGCTTCAATTCAGCTCCGGTTCCTCAACGATCAAGGCG	603
QY	3061	AGTTACATGATCCCCATGTTGTGCAAAAAGCGGTTAGCTCTTCGGTCTCTCCGATCGT	3120
Db	602	AGTTACATGATCCCCATGTTGTGCAAAAAGCGGTTAGCTCTTCGGTCTCTCCGATCGT	543
QY	3121	TGTCAGAGTAAGTTGGCGGAGTGTATCACTCATGTTTATGGCAGCAGTGCATAATTC	3180
Db	542	TGTCAGAGTAAGTTGGCGGAGTGTATCACTCATGTTTATGGCAGCAGTGCATAATTC	483
QY	3181	TCCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGACTGTGTAGTACTCAACCAAGTC	3240
Db	482	TCCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGACTGTGTAGTACTCAACCAAGTC	423
QY	3241	ATTTCTGAGAATAGTGTATGCGCGACCGAGTTGCTCTTTGCCGGGGTCAATAACGGGATAA	3300
Db	422	ATTTCTGAGAATAGTGTATGCGCGACCGAGTTGCTCTTTGCCGGGGTCAATAACGGGATAA	363
QY	3301	TAACCGCGGCACATAGCAGNACTTTAAAGTGCTCATCATTTGAAAAACGTTCTTCGGGGG	3360
Db	362	TAACCGCGGCACATAGCAGNACTTTAAAGTGCTCATCATTTGAAAAACGTTCTTCGGGGG	303
QY	3361	AAAACTCTCAAGGATCTTACCCGCTGTGAGATCCAGTTTCGATTAACCCACTCGTGACCC	3420
Db	302	AAAACTCTCAAGGATCTTACCCGCTGTGAGATCCAGTTTCGATTAACCCACTCGTGACCC	243
QY	3421	CAACTGATCTTCAGCATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAAACAGGAAG	3480
Db	242	CAACTGATCTTCAGCATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAAACAGGAAG	183
QY	3481	GCAAAATGCGCAAAAAGGGAATTAAGGCGACAGGAAATTTGAATACTCATACTCTT	3540
Db	182	GCAAAATGCGCAAAAAGGGAATTAAGGCGACAGGAAATTTGAATACTCATACTCTT	123
QY	3541	CTTTTTCATATTTATGAAGCATTTATTCAGGTTTATTTGCTCATGAGCGGATACATATT	3600
Db	122	CTTTTTCATATTTATGAAGCATTTATTCAGGTTTATTTGCTCATGAGCGGATACATATT	63
QY	3601	TGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGGCGACATTTTCCCGCAAAAGTGCC	3660
Db	62	TGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGGCGACATTTTCCCGCAAAAGTGCC	3
QY	3661	AC 3662	
Db	2	AC 1	

RESULT 7
ADM68448/c
ID ADM68448 standard; DNA; 3637 BP.
XX

AC	ADM68448;	
XX	03-JUN-2004 (first entry)	
DE	Cycle 3 green fluorescent protein, GFP, DNA construct.	
XX	ds; mismatch endonuclease; endonuclease; gene shuffling technology;	
KW	single nucleotide polymorphism; cancer susceptibility;	
KW	sequence variation redistribution; cycle 3 GFP;	
KW	green fluorescent protein.	
XX	Aequorea victoria.	
OS	Synthetic.	
XX	US2003157682-A1.	
PN	21-AUG-2003.	
PD	31-JAN-2003; 2003US-00356708.	
PF	01-FEB-2002; 2002US-0353722P.	
XX	14-MAR-2002; 2002US-00098155.	
PR	01-AUG-2002; 2002US-00211079.	
XX	(PADG/) PADGETT H S.	
PA	(VAEW/) VAEWHONGS A A.	
PA	(VOJD/) VOJDANI F S.	
PA	(SMIT/) SMITH M L.	
PA	(LIND/) LINDBO J A.	
PA	(FITZ/) FITZMAURICE W P.	
XX	Padgett HS, Vaewhongs AA, Vojdani FS, Smith ML, Lindbo JA; Fitzmaurice WP;	
PI	WPI; 2003-766176/72.	
XX	Making a mismatch endonuclease, useful in gene shuffling and in detection of single nucleotide polymorphisms, comprises transfecting a host with a recombinant viral vector including a polynucleotide encoding a mismatch endonuclease.	
PT	Example 9; SEQ ID NO 17; 79pp; English.	
PS	The invention relates to a method of making a mismatch endonuclease enzyme comprising transfecting a host plant, animal, yeast, fungus or bacterium with a recombinant viral vector that encodes a polynucleotide sequence for a mismatch endonuclease, growing the host so that the polynucleotide is expressed, and extracting the mismatch endonuclease enzyme from the host. The method is useful for making mismatch endonuclease enzymes, for obtaining peptides and polynucleotides with desired functional properties and for detecting mutations. The mismatch endonuclease enzymes are useful in gene shuffling technology for developing new genes, in detecting single nucleotide polymorphisms for e.g. detecting evidence of cancer susceptibility, or in redistributing sequence variations between non-identical polynucleotide sequences. The present sequence represents the cycle 3 green fluorescent protein, GFP, DNA construct.	
XX	Sequence 3637 BP; 942 A; 892 C; 870 G; 933 T; 0 U; 0 Other;	
SQ	Query Match 72.8%; Score 2664.8; DB 11; Length 3637; Best Local Similarity 83.8%; Pred. No. 0; Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;	
QY	1 CTAATTCGTAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60	
Db	3637 CTAATTCGTAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 3578	
QY	61 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATATAATCAAAGATAGACCGA 120	
Db	3577 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATATAATCAAAGATAGACCGA 3518	
QY	121 GATAGGGTTGAGTGTGTTCAGTTTGGAAACAAGAGTCCACTATTATAAAGACGTGGACTC 180	

Db 3517 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAGAAACGTGGACTC 3458
QY 181 CAAAGCTCAAAGGGCGAAACCCGCTTATCAGGGCGATGGCCCACTACGTGAACCATCAC 240
Db 3457 CAAAGCTCAAAGGGCGAAACCCGCTTATCAGGGCGATGGCCCACTACGTGAACCATCAC 3398
QY 241 CTAATCAAGTTTTTTGGGTGAGGTGCGGTAAAGCACTAAATCGGAACCCCTAAAGGAG 300
Db 3397 CTAATCAAGTTTTTTGGGTGAGGTGCGGTAAAGCACTAAATCGGAACCCCTAAAGGAG 3338
QY 301 CCCCCCATTTAGAGCTTGACGGGGAAGCCGCGAAGCTGGCGAGAAAGGAAGAA 360
Db 3337 CCCCCCATTTAGAGCTTGACGGGGAAGCCGCGAAGCTGGCGAGAAAGGAAGAA 3278
QY 361 AGCGAAGGAGCGGGCGCTAGGCGCTGGCAAGTGTAGCGGTCACTGCGCGTAAACCAC 420
Db 3277 AGCGAAGGAGCGGGCGCTAGGCGCTGGCAAGTGTAGCGGTCACTGCGCGTAAACCAC 3218
QY 421 CACACCGCGCGCTTAATGCGCGCTCAGAGGCGGTGCCATTCGCCATTCAGGCTGCG 480
Db 3217 CACACCGCGCGCTTAATGCGCGCTCAGAGGCGGTGCCATTCGCCATTCAGGCTGCG 3158
QY 481 CAACTGTTGGGAAGGGCGATCGGTGCGGCGCTCTTCGCTATTACGCCAGCTGGCGAAG 540
Db 3157 CAACTGTTGGGAAGGGCGATCGGTGCGGCGCTCTTCGCTATTACGCCAGCTGGCGAAG 3098
QY 541 GGGATGCTGCAAGCGGTAAGTTGGGTAAAGCCAGCGGTTTCCAGTACAGAGCTTG 600
Db 3097 GGGATGCTGCAAGCGGTAAGTTGGGTAAAGCCAGCGGTTTCCAGTACAGAGCTTG 3038
QY 601 TAAACGACGGCGAGTGGCGGCTCGTTCAFTCACTGTTTTGAAACCCGTGGAGGACGG 660
Db 3037 TAAACGACGGCGAGTGGCGGCTAAAGTACGACTCACTATAGGGCGAATTTGGAGTCCA 2978
QY 661 GCAGACTCGGCGTGAATGTGTTTACAGCGTATGGAGCAGATGAAGTCTCGACAC 720
Db 2977 CCGCGGTGGCGG-----CCGCTCTAGAACTAGTGGATCCCC--CGG 2939
QY 721 GCTGCAGAAACACGACGTAGATTAAACCTAGAAAGATAATCATATTGTGACGTCAGGTAA 780
Db 2938 GCTGCAGGAATCTTATTATTGTATGTATTCATCCATGCCATGTGTAATCCGACGAGTTA 2879
QY 781 AGATAATCATCGGTAAATTTGACGATGGGATCTGTAATACGACTCACTATAGGGCGAAT 840
Db 2878 CAAACTCAAGAGGACCATGTGTCACGCTTTTCGTTGGGATCTTTCGAAAGGGCAGATT 2819
QY 841 TGGGTACCGGGCCCCCTCGAGTCCAGGTCCAGGTATCGAATAGCTTGATATCGAATTCCTGC 900
Db 2818 GTGTCACAGGTAAATGGTTGTCTGTTAAAGGACAGGGCCATCGCCAAATTTGAGTATTTT 2759
QY 901 AGCCCGGGGATCCACTAGTTCTPAGAGCGCGCCACCGCGGTGGAGCTCCAGCTTTTGT 960
Db 2758 GTTGATAATGGTCTGCTAGTTGAACGATCCATCTTCAATGTTGTGGCGAATTTGAAAT 2699
QY 961 TCCCTTTAGTGGGTTAATTAGATCCCATCGGTCAATTTTACGACAGACTATCTTTCTAG 1020
Db 2698 TAGCTTTGATTCATCTTTTGTGTCGCGTGATGATACATTTGTGTGAG-TTATAG 2640
QY 1021 GGTAACTAGCTGCAATCAGGATCATATCGTGGGCTTTTTTCCGGCTCAGTCACTGCGC 1080
Db 2639 TTGTACTCAGATTGTGTCGAGAAATTTTCCATCTTCTTTAAATCAATACCTTTTAAAC 2580
QY 1081 CAAAGTGGCGCTATCTGGGCATCGGGGAGGAAGCCCGTGCTTTTCCCGAGGTTG 1140
Db 2579 TCGATACGATTAACAAGGGTATCACTTTCAAACTTGACTTCAGCAGCGGTCTTGTTAGTTC 2520
QY 1141 AAGCGGCATGGAAGAGTTTGGCGAGGATGACTGCTGCTGCAATTTGACGTTGAGCGGAAAC 1200
Db 2519 CCGTCATCTTTGAAGATATAGTGCCTTCTGTACATAACCTTCGGGCATGGCACTTTG 2460
QY 1201 GCAGGTTTACCATGATTTGGGAAGGTGTGGCCATGACGCGCTTTAAACGGTGAATCTGT 1260

Db 2459 AAAAAGTCATGCGGTTTTCATATGATCCGATTAACGGGAAAAGCAATTGAACACCATAAAG 2400
QY 1261 TCGTTCAGCCACCTGGGATACAGTTCTGTCGGGCTTTTCCGGACACACAGTTCCGGATGG 1320
Db 2399 AAAGTAGTGACAAGTGTGGCCATGGAAACAGGTAGTTTTTCCAGTAGTGCAAAATAAATTTA 2340
QY 1321 TCAGCCCGAAGCGCATCAGCAACCCGAAACAAATACCGGCGACAGCCGGAACCTGCGGTGCG 1380
Db 2339 AGGGTAAGCTTTTCCGTATGTAGCATCACCTTCACTCTCCACTGACAGAAATTTGTGTC 2280
QY 1381 GTGTGCAGATTAATGACAGCGGTGCGGCTGGGATATTACGTGACGAGGACGGGTATC 1440
Db 2279 CCATTAACATCACCATCTTAATTCACAAGAAATTTGGACAACCTCCAGTGAAGTTCTTCT 2220
QY 1441 CTGGCTGGATCCGCAAGAAATGGAATGACATGCCGTGAGTTA CCGCGCGGCGCGCTT 1500
Db 2219 C---CTTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTAATTGCGGCTT 2163
QY 1501 GGGTAAATCATGTCATAGCTGTTTCTGTGTGAATTTGTTATCTCCGCTCACAAATTCACA 1560
Db 2162 GGGTAAATCATGTCATAGCTGTTTCTGTGTGAATTTGTTATCTCCGCTCACAAATTCACA 2103
QY 1561 CAACTACGAGCCGGAAGCATAAAGTGTAAAGCTTGGGTGCTTAATGAGTGAAGTAACT 1620
Db 2102 CAACTACGAGCCGGAAGCATAAAGTGTAAAGCTTGGGTGCTTAATGAGTGAAGTAACT 2043
QY 1621 CACATTAATTTGGGTGCGCTCACTGCCCTTTTCACTGCGGGAACCTGTGTCGCGAGCT 1680
Db 2042 CACATTAATTTGGGTGCGCTCACTGCCCTTTTCACTGCGGGAACCTGTGTCGCGAGCT 1983
QY 1681 GAATTAATGAATCGGCAACCGCGGGGAGAGCGGTTTTCGTTATTTGGGCGCTCTTCCGC 1740
Db 1982 GCATTAATGAATCGGCAACCGCGGGGAGAGCGGTTTTCGTTATTTGGGCGCTCTTCCGC 1923
QY 1741 TTCTCGCTCACTGACTCGCTCGCTCGGTTCGTCGTCGCGTGGCGAGCGTATCAGTCA 1800
Db 1922 TTCTCGCTCACTGACTCGCTCGCTCGGTTCGTCGTCGCGTGGCGAGCGGATCAGTCA 1863
QY 1801 CTCAAAGGCGGTAATACGGTTATCCAAGAAATCAGGGGATAACGCAAGGAAAGAACTGTG 1860
Db 1862 CTCAAAGGCGGTAATACGGTTATCCAAGAAATCAGGGGATAACGCAAGGAAAGAACTGTG 1803
QY 1861 AGCAAAGGCGGAGCAAAAGGCCAGGAACCGTAAAGAGCGCGTTCCTGCGCTTTTTCCTCA 1920
Db 1802 AGCAAAGGCGGAGCAAAAGGCCAGGAACCGTAAAGAGCGCGTTCCTGCGCTTTTTCCTCA 1743
QY 1921 TAGGCTCGGCCCCCTGACGAGCATCACAAAATCGAACCTCAAGTCAAGAGTGGCGAAA 1980
Db 1742 TAGGCTCGGCCCCCTGACGAGCATCACAAAATCGAACCTCAAGTCAAGAGTGGCGAAA 1883
QY 1981 CCGCAGAGGACTATAAAGATACAGCGGTTTTCCCTCGGAAGCTCCCTCGTGGCTCTCC 2040
Db 1682 CCGCAGAGGACTATAAAGATACAGCGGTTTTCCCTCGGAAGCTCCCTCGTGGCTCTCC 1623
QY 2041 TGTTCGACCTTCGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGC 2100
Db 1622 TGTTCGACCTTCGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGC 1563
QY 2101 GCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTTCGGGTAGTCTCGTTCGCTCCAAGCT 2160
Db 1562 GCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTTCGGGTAGTCTCGTTCGCTCCAAGCT 1503
QY 2161 GGGCTGTGTGACGAACCCCGCTTCAGCGCCGACCGCTGCGCTTATCCGGTAACTATCG 2220
Db 1502 GGGCTGTGTGACGAACCCCGCTTCAGCGCCGACCGCTGCGCTTATCCGGTAACTATCG 1443
QY 2221 TCTTGAGTCCAAACCGGGTAAGACAGCTTATCGCCACTGGCGAGCAGCACTGGTAAACAG 2280
Db 1442 TCTTGAGTCCAAACCGGGTAAGACAGCTTATCGCCACTGGCGAGCAGCACTGGTAAACAG 1383
QY 2281 GATTAGCAGAGCGGTATGTAGGGGTGCTACAGAGTTCTTCAAGTGGTGGCTAACTA 2340
Db 1382 GATTAGCAGAGCGGTATGTAGGGGTGCTACAGAGTTCTTCAAGTGGTGGCTAACTA 1323

QY 2341 CGGCTACACTAGAGCAGATATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 2400
DB |||||
DB 1322 CGGCTACACTAGAGCAGATATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 1263
QY 2401 AAAAGAGTTGGTAGCTCTTGATTCGGCAAAACAAACACCGCTGGTAGCGGTGTTTTT 2460
DB |||||
DB 1262 AAAAGAGTTGGTAGCTCTTGATTCGGCAAAACAAACACCGCTGGTAGCGGTGTTTTT 1203
QY 2461 TCTTTGCAAGCAGCAGATTTACCGCAGAAAAAGGATCTCAAGAGATCTCTTTGATCTT 2520
DB |||||
DB 1202 TCTTTGCAAGCAGCAGATTTACCGCAGAAAAAGGATCTCAAGAGATCTCTTTGATCTT 1143
QY 2521 TTCTACCGGGTCTGACGCTCAGTGGAAACGAAATCTCACGTTAAGGATTTTGGTCATGAG 2580
DB |||||
DB 1142 TTCTACCGGGTCTGACGCTCAGTGGAAACGAAATCTCACGTTAAGGATTTTGGTCATGAG 1083
QY 2581 ATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCAAT 2640
DB |||||
DB 1082 ATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCAAT 1023
QY 2641 CTAAAGTATATAGTAAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAAGCACC 2700
DB |||||
DB 1022 CTAAAGTATATAGTAAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAAGCACC 963
QY 2701 TATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGACTCCCGTCTGTAGAT 2760
DB |||||
DB 962 TATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGACTCCCGTCTGTAGAT 903
QY 2761 AACTACGATACCGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCCGAGACCC 2820
DB |||||
DB 902 AACTACGATACCGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCCGAGACCC 843
QY 2821 ACGCTCACCGGCTCCAGATTTATCAGCAATAACAGCAGCCGGAAGGGCGGCGAG 2880
DB |||||
DB 842 ACGCTCACCGGCTCCAGATTTATCAGCAATAACAGCAGCCGGAAGGGCGGCGAG 783
QY 2881 AAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTATTTGTCGGGGAAGCTAG 2940
DB |||||
DB 782 AAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTATTTGTCGGGGAAGCTAG 723
QY 2941 AGTAAGTAGTTCCGCAAGTAATAGTTTGGCAACGGTTGTGTCATTTGTCAGGCAATCGT 3000
DB |||||
DB 722 AGTAAGTAGTTCCGCAAGTAATAGTTTGGCAACGGTTGTGTCATTTGTCAGGCAATCGT 663
QY 3001 GGTGTACGCTCGTCTGTTGGTATGGCTTCATTCAGCTCCGGTTCCTCAACGATCAAGCG 3060
DB |||||
DB 662 GGTGTACGCTCGTCTGTTGGTATGGCTTCATTCAGCTCCGGTTCCTCAACGATCAAGCG 603
QY 3061 AGTTACATGATCCCCATGTTGTGCAAAAAAGCGTTAGTCTCTTCGGTCTCTCCGATCGT 3120
DB |||||
DB 602 AGTTACATGATCCCCATGTTGTGCAAAAAAGCGTTAGTCTCTTCGGTCTCCGATCGT 543
QY 3121 TGTCAAGAGTAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATTC 3180
DB |||||
DB 542 TGTCAAGAGTAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATTC 483
QY 3181 TCTTACTGTATGCCATCGTAAAGTGTCTTTCTGTGACTGGTGAGTACTCAACCAATC 3240
DB |||||
DB 482 TCTTACTGTATGCCATCGTAAAGTGTCTTTCTGTGACTGGTGAGTACTCAACCAATC 423
QY 3241 ATTCTGAGAATAGTGTATGCGCGCAGCCAGTTGCTCTTCCCGCGGTCAATACGGGATAA 3300
DB |||||
DB 422 ATTCTGAGNAATAGTGTATGCGCGCAGCCAGTTGCTCTTCCCGCGGTCAATACGGGATAA 363
QY 3301 TACCGCGCACATAGCAGAACTTTTAAAGTGTCTCATATGTGAAAAACGTTCTTCGGGGCG 3360
DB |||||
DB 362 TACCGCGCACATAGCAGAACTTTTAAAGTGTCTCATATGTGAAAAACGTTCTTCGGGGCG 303
QY 3361 AAACTCTCAAGGATCTTACCGCTGTGAGTCCAGTTCGATGAACCACTCGTGCACC 3420
DB |||||
DB 302 AAACTCTCAAGGATCTTACCGCTGTGAGTCCAGTTCGATGAACCACTCGTGCACC 243

RESULT 8

ADP26601/c

ID ADP26601 standard; DNA; 3637 BP.

XX

AC ADP26601;

XX

DT 26-AUG-2004 (first entry)

XX

DE Green fluorescent protein (GFP) plasmid DNA #2.

XX

KW Sequence variation; heteroduplex; transcription; DNA integration;

KW ribozyme expression; gene; ds; green fluorescent protein; GFP.

XX

OS Aequorea victoria.

OS Synthetic.

XX

PN US2004110130-A1.

XX

PD 10-JUN-2004.

XX

PF 25-OCT-2002; 2002US-00280913.

XX

PR 02-FEB-2001; 2001US-0266386P.

PR

PR 14-FEB-2001; 2001US-0268785P.

PR

PR 01-FEB-2002; 2002US-00066390.

PR

PR 08-AUG-2002; 2002US-0402342P.

XX

(LARG-) LARGE SCALE BIOLOGY CORP.

PA

Padgett HS, Lindbo JA, Fitzmaurice WP;

XX

DR WPI; 2004-440326/41.

XX

XX

PS Example 5; SEQ ID NO 17; 75pp; English.

XX

The invention relates to an in vitro method of redistributing sequence variations between non-identical polynucleotide sequences, comprising making a heteroduplex polynucleotide from two non-identical polynucleotides, introducing a nick in the second strand at or near a base pair mismatch site, removing the mismatched base(s) from the mismatch site where the nick occurred and using the first strand as a template to replace the removed base(s) with bases that complement the base(s) in the first strand. The invention also relates to an in vitro method of making a population of sequence variants from a heteroduplex polynucleotide sequence, obtaining a polynucleotide sequence encoding a desired functional property and identifying a reassorted DNA molecule encoding a protein with a desired functional property. The method is useful for generating an improved polynucleotide sequence or a population

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC	of improved polynucleotide sequences possessing at least one desired									
CC	phenotypic characteristic (e.g., promotes transcription of linked									
CC	polynucleotides), where such polynucleotides are useful for expression									
CC	from a plant, animal, fungal, yeast, or bacterial expression vector, for									
CC	integration to form a transgenic plant, animal or microorganism, and for									
CC	expression of a ribozyme. This sequence represents DNA used in the scope									
CC	of the invention.									
XX										
SQ	Sequence 3637 BP; 942 A; 892 C; 870 G; 933 T; 0 U; 0 Other;									
	Query Match	72.8%;	Score 2664.8;	DB 12;	Length 3637;					
	Best Local Similarity	83.8%;	Pred. No. 0;							
	Matches 3070;	Conservative 0;	Mismatches 567;	Indels 25;	Gaps 4;					
Qy	1	CTAAATTGTAAGCGTTAAATATTTTGTGTTAAATTCGCGTTAAATTTTGTGTTAAATCAGCTC	60							
Db	3637	CTAAATTGTAAGCGTTAAATATTTTGTGTTAAATTCGCGTTAAATTTTGTGTTAAATCAGCTC	3578							
Qy	61	ATTTTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATAAATCAAAAGATAGACCGA	120							
Db	3577	ATTTTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATAAATCAAAAGATAGACCGA	3518							
Qy	121	GATAGGTTGAGTGTGTTCCAGTTTTCGAAACAGAGTCCACTATTAAAGAACTGGGACTC	180							
Db	3517	GATAGGTTGAGTGTGTTCCAGTTTTCGAAACAGAGTCCACTATTAAAGAACTGGGACTC	3458							
Qy	181	CAACGTCAAAAGGCGAAAAACCGTCTATCAGGCGGATGGCCCACTACGTGAACCATCACC	240							
Db	3457	CAACGTCAAAAGGCGAAAAACCGTCTATCAGGCGGATGGCCCACTACGTGAACCATCACC	3398							
Qy	241	CTAATCAAGTTTTTTTGGGTCGAGGTGCGGTAAAGCACTAAATCGGAACCCCTAAAGGGAG	300							
Db	3397	CTAATCAAGTTTTTTTGGGTCGAGGTGCGGTAAAGCACTAAATCGGAACCCCTAAAGGGAG	3338							
Qy	301	CCCCGATTTAGAGCTTTGACGGGAAAGCCGCGAACGTGGCGAGAAAGCAAGGAGAA	360							
Db	3337	CCCCGATTTAGAGCTTTGACGGGAAAGCCGCGAACGTGGCGAGAAAGCAAGGAGAA	3278							
Qy	361	AGCGAAAGGAGCGGCGCTAGGCGCGCTGGCAAGTGTAGCGGTACACGTGCGCGGTAAACAC	420							
Db	3277	AGCGAAAGGAGCGGCGCTAGGCGCGCTGGCAAGTGTAGCGGTACACGTGCGCGGTAAACAC	3218							
Qy	421	CACACCCGCGGCTTAATGCGCGCTACAGGCGCGTCCCAATTCGCCATTCAGGCTGCG	480							
Db	3217	CACACCCGCGGCTTAATGCGCGCTACAGGCGCGTCCCAATTCGCCATTCAGGCTGCG	3158							
Qy	481	CAACTGTTGGGAAGGCGGATCGGTGCGGCGCTCTTCGCTATTACGCCAGCTGCGGAAAGG	540							
Db	3157	CAACTGTTGGGAAGGCGGATCGGTGCGGCGCTCTTCGCTATTACGCCAGCTGCGGAAAGG	3098							
Qy	541	GGGATGTGTCGAAGGCGATTAAAGTTGGGTAAACGCCAGGCTTTTCCAGTCAACGAGCTTG	600							
Db	3097	GGGATGTGTCGAAGGCGATTAAAGTTGGGTAAACGCCAGGCTTTTCCAGTCAACGAGCTTG	3038							
Qy	601	TAAACGACGCGCCAGTGAAGCGCGCTCGTTCAATTCACGTTTTTTTGAACCCGCTGGAGACGG	660							
Db	3037	TAAACGACGCGCCAGTGAAGCGCGCTTAATACGACTCACTATAGGGCGAATTGGAGCTCCA	2978							
Qy	661	GCAGACTCGCGGTGCAAAATGTGTTTTACAGGTGATGAGCAGATGAAGATGCTCGACAC	720							
Db	2977	CCGCGTGGCGG-----CCGCTCTAGAACCTAGTGGATCCCC--CGG	2939							
Qy	721	GCTGCAGAACACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTACGTTAA	780							
Db	2938	GCTGCAGGAATTTCTATTGTTATAGTTTCATCCATGCCATGTGTAATCCCGAGCAGTGTTA	2879							
Qy	781	AGATAATCATGCGTAAATTTGACGCATGGGATCTGTAAATACGACTCACTATAGGGCGAAT	840							
Db	2878	CAAACTCAAGAAGGACCATGTGTCACGCTTTTCGTTGGGATCTTTTCGAAGGGCAGATT	2819							
Qy	841	TGGGTACCGGGCCCCCTCGAGGTGCAGCGGTATCGATAAGCTTGATATCGAATTCCTGCG	900							
Db	2818	GTGTCGACAGGTAATGGTTGTCTCGTTAAAGGACAGGGCCATCGCCAAATTTGGAGTATTTT	2759							

QY	901	AGCCCCGGGGATCCACTAGTTCTTAGAGCGCGCGCCACCCGCGTGGAGTCCAGCTTTTGT	960						
DB	2758	GTTGATAATGGTCTGCTAGTTGCAACGGATCCATCTTCAATGTTGTGGCGAAATTTGAAGT	2699						
QY	961	TCCCTTTAGTGAGGGTTAATTAGATCCCATGGTCAATTTTACGCAGACTATCTTCTAG	1020						
DB	2698	TAGCTTTGATTCATTTCTTTTGTCTGCGGTGATGTATACATTTGTGTAG-TTATAG	2640						
QY	1021	GGTTAATCTAGCTGCATCAGGATCATATCGTGGGCTCTTTTTCGGCTCAGTCATCGCC	1080						
DB	2639	TTGTACTCGAGTTTGTGTCGAGAAATGTTTCCATCTTCTTTAAATCAATACCTTTTAAC	2580						
QY	1081	CAAGCTGGCGCTATCTGGGCATCGGGAGGAAGAACCCGCTGCTTTTCCCGGAGGTG	1140						
DB	2579	TCGATACGATTAAACAAGGGTATCACCTTCAAACTTGACTTCAGCACGCGCTTGTAGTTC	2520						
QY	1141	AAGCGGCATGGAAGAGTTTTCGCGAGGATGACTGCTGCTGCAATGACGTTTGAGCGAAAC	1200						
DB	2519	CGGTCACTTTTGAAGAATATAGTGGCTTCTGTACATAAACCTTCGGGCATGGCACTCTTG	2460						
QY	1201	GCACGTTTACCATGATGATTTCGGGAAGGTGTGGCCATGCAACGCTTTTAAACGGTGAAC	1260						
DB	2459	AAAAAGTCATCGCTTTCATATGATCCGATTAACGGGAAAGCAATTGACACCATTAAGAG	2400						
QY	1261	TCGTTACGCCACCTGGGATACAGTTCTGTCGGCTTTTTCGGGACACAGTTCGCGATGG	1320						
DB	2399	AAAGTAGTGACAAGTGTGGCCATGGAAACAGGTAGTTTTCAGTAGTAGTCAAAATAAATTA	2340						
QY	1321	TCAGCCCGAAGCGCATCAGCAACCCGACAAATACCGGCGACAGCCGGAACCTGCCGTGCG	1380						
DB	2339	AGGGTAAGCTTTCCGTATGTAGCATACCTTCACTCCCTCTCCACTGACAGAAATTTGTGC	2280						
QY	1381	GTGTGCAGATTAAATGACAGCGGTGGCGCTGGGATATTACGTGACGAGGAGCGGTATC	1440						
DB	2279	CCATTAACTACCATCTAATTTCAACAGAAATTGGGCAACTCCAGTGAAGTTCTTCT	2220						
QY	1441	CTGGCTGGATCCGCGAGAAATGGAATGATGATACCGGTGAGTTACCCGCGCGCGCGCTT	1500						
DB	2219	C---CTTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAGGTTAATTTGCGGCTT	2163						
QY	1501	GGCGTAATCATGGTCTAGCTGTTCTGTGTGTAATTTGTTATCCGCTCACAAATTTCCCA	1560						
DB	2162	GGCGTAATCATGGTCTAGCTGTTCTGTGTGTAATTTGTTATCCGCTCACAAATTTCCCA	2103						
QY	1561	CAACATACGAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTAATAGTGAGCTAACT	1620						
DB	2102	CAACATACGAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTAATAGTGAGCTAACT	2043						
QY	1621	CACATTAATTGCTTTCGCTCATCTCCGCTTTTCCAGTCGGGAAACCTGTCGTGCGAGCT	1680						
DB	2042	CACATTAATTGCTTTCGCTCATCTCCGCTTTTCCAGTCGGGAAACCTGTCGTGCGAGCT	1983						
QY	1681	GCATTAATGAATCGGCAACCGCGGGAGAGCGGTTTTCGTATTTGGCGCTTCTCCGC	1740						
DB	1982	GCATTAATGAATCGGCAACCGCGGGAGAGCGGTTTTCGTATTTGGCGCTTCTCCGC	1923						
QY	1741	TTCTCTCGCTCACTGACTCGCTCGCTCGGTTCGCTTCGCTGCGCGAGCGGTATCAGCTCA	1800						
DB	1922	TTCTCTCGCTCACTGACTCGCTCGCTCGGTTCGCTTCGCTGCGCGAGCGGTATCAGCTCA	1863						
QY	1801	CTCAAAGCGGTAATACGGTTATCCAGAAATCAGGGGATAACGCGAGGAAAGAACATGTG	1860						
DB	1862	CTCAAAGCGGTAATACGGTTATCCAGAAATCAGGGGATAACGCGAGGAAAGAACATGTG	1803						
QY	1861	ACCAAAGCCAGCAAAAGGCCAGGAAACGTTAAAGGCGCGGTTGCTGCGCTTTTCCA	1920						
DB	1802	AGCAAAGCCAGCAAAAGGCCAGGAAACGTTAAAGGCGCGGTTGCTGCGCTTTTCCA	1743						
QY	1921	TAGGCTCGCCCCCTTGACGAGCATCAAAAATCGACCTCAAGTCAGAGGTGGCGAAA	1980						
DB	1742	TAGGCTCGCCCCCTTGACGAGCATCAAAAATCGACCTCAAGTCAGAGGTGGCGAAA	1683						

1981 CCCGACGAGCTATAAAGATACACGCGTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCC 2040
1682 CCCGACGAGCTATAAAGATACACGCGTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCC 1623
2041 TGTTCGACCCCTGCGGCTTACCGGATACCTGTCGCGCTTTTCCCTTCCGGGAAGCGTGGC 2100
1622 TGTTCGACCCCTGCGGCTTACCGGATACCTGTCGCGCTTTTCCCTTCCGGGAAGCGTGGC 1563
2101 GCTTTCTCATAGCTCACCGTGTAGGTATCTCAGTTTCGGTGTAGGTGCTTCGCTCAAGCT 2160
1562 GCTTTCTCATAGCTCACCGTGTAGGTATCTCAGTTTCGGTGTAGGTGCTTCGCTCAAGCT 1503
2161 GGGCTGTGTGACGAAACCCCGCTTACGCGCCGACCGCTGCGCTTATCCGGTAACTATCG 2220
1502 GGGCTGTGTGACGAAACCCCGCTTACGCGCCGACCGCTGCGCTTATCCGGTAACTATCG 1443
2221 TCTTGAGTCCAAACCCGGTAAAGACACGACTTATCGGCACCTGGCAGCAGCACTGGTAAACAG 2280
1442 TCTTGAGTCCAAACCCGGTAAAGACACGACTTATCGGCACCTGGCAGCAGCACTGGTAAACAG 1383
2281 GATTAGCAGCGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCCTAACTA 2340
1382 GATTAGCAGCGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCCTAACTA 1323
2341 CGGCTACACTAAGAGCAGATATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTTACCTTCGG 2400
1322 CGGCTACACTAAGAGCAGATATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTTACCTTCGG 1263
2401 AAAAAAGTGTAGTCTTGTATCCGGCAAAACAAACCCAGCGTGTGAGCGGTGTTTTT 2460
1262 AAAAAAGTGTAGTCTTGTATCCGGCAAAACAAACCCAGCGTGTGAGCGGTGTTTTT 1203
2461 TGTTTTGCAAGCAGAGTATTCGCGCAGAAAAAGGATCTCAAGAGATCCTTTTCATCTT 2520
1202 TGTTTTGCAAGCAGAGTATTCGCGCAGAAAAAGGATCTCAAGAGATCCTTTTCATCTT 1143
2521 TTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAAGGATTTTGGTCATGAG 2580
1142 TTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAAGGATTTTGGTCATGAG 1083
2581 ATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAAATGAAGTTTAAATCAAT 2640
1082 ATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAAATGAAGTTTAAATCAAT 1023
2641 CTAAAGTATATAGTAAACTTGTCTGACAGTTTACCAATGCTTAAATCAGTGAGGCACC 2700
1022 CTAAAGTATATAGTAAACTTGTCTGACAGTTTACCAATGCTTAAATCAGTGAGGCACC 963
2701 TATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGAT 2760
962 TATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGAT 903
2761 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCAGACCC 2820
902 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCAGACCC 843
2821 AGCTCAGCGGCTCAGATTTATACGATATAACAGCCAGCCGGAAGCGCGAGCGAG 2880
842 AGCTCAGCGGCTCAGATTTATACGATATAACAGCCAGCCGGAAGCGCGAGCGAG 783
2881 AAGTGGTCTGCAAACTTTATCCGCTCCATCCAGTCTATTATTTGTCGGGGAAGCTAG 2940
782 AAGTGGTCTGCAAACTTTATCCGCTCCATCCAGTCTATTATTTGTCGGGGAAGCTAG 723
2941 AGTAAGTATGTCGAGTTTAAATAGTTTGGCAGAACTGTTGTCATTTGCTACAGGCATCGT 3000
722 AGTAAGTATGTCGAGTTTAAATAGTTTGGCAGAACTGTTGTCATTTGCTACAGGCATCGT 663
3001 GGTGTACGCTCGTGTGTTGATGGTTCATTCAGCTCCGTTCCCAACGATCAAGGG 3060
662 GGTGTACGCTCGTGTGTTGATGGTTCATTCAGCTCCGTTCCCAACGATCAAGGG 603
3061 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGTCTCTTTCGCTCCTCCGATCGT 3120

602 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGTCTCTTCGTTCTCCGATCGT 543
3121 TGTCAAGATAGTTTGGCGCAGTGTATCACTCATGTTTATGGCAGCACTGCATAATTC 3180
542 TGTCAAGATAGTTTGGCGCAGTGTATCACTCATGTTTATGGCAGCACTGCATAATTC 483
3181 TCTTACTGTCTATGCGCATCCGTAAGATGCTTTTCTGTGATGGTGTAGTACTCAACCAAGTC 3240
482 TCTTACTGTCTATGCGCATCCGTAAGATGCTTTTCTGTGATGGTGTAGTACTCAACCAAGTC 423
3241 ATTCTGAGAATAGTGTATGCGCGCAGCCGAGTTGCTTTCGCCGGCGTCAATAACGGGATAA 3300
422 ATTCTGAGAATAGTGTATGCGCGCAGCCGAGTTGCTTTCGCCGGCGTCAATAACGGGATAA 363
3301 TACCGCGCCACATAGCAGAACTTTAAAAGTGTCTATCAATTTGGAACAACTTCTTCGGGG 3360
362 TACCGCGCCACATAGCAGAACTTTAAAAGTGTCTATCAATTTGGAACAACTTCTTCGGGG 303
3361 AAACTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCGATGTAACCCACTCGTGCACC 3420
302 AAACTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCGATGTAACCCACTCGTGCACC 243
3421 CAACTGTATCTTTCAGCATCTTTTACTTTTCCAGCGCTTTTCTGGTGAGCAAAAACAGGAAG 3480
242 CAACTGTATCTTTCAGCATCTTTTACTTTTCCAGCGCTTTTCTGGTGAGCAAAAACAGGAAG 183
3481 GCAAAATCCGCAAAAAGGGAATTAAGGCGCAGCAAGAAATGTTGAATACTCATACTCTT 3540
182 GCAAAATCCGCAAAAAGGGAATTAAGGCGCAGCAAGAAATGTTGAATACTCATACTCTT 123
3541 CCTTTTCAATATTTTGAAGCATTTTATCAGGTTTATTTGTTCTCATGAGCGGATACATATT 3600
122 CCTTTTCAATATTTTGAAGCATTTTATCAGGTTTATTTGTTCTCATGAGCGGATACATATT 63
3601 TCAATGTTATTTAGAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGCAAAAGTGCC 3660
62 TGAATGTTATTTAGAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGCAAAAGTGCC 3
3661 AC 3662
2 AC 1

RESULT 9
ADQ88423/c
ID ADQ88423 standard; DNA; 3637 BP.
XX
AC ADQ88423;
XX
DT 18-NOV-2004 (first entry)
XX
DE PBSC3GFP DNA encoding Jellyfish cycle 3 GFP protein.
XX
KW Molecular biology; jellyfish; green fluorescent protein; GFP; ds.
XX
OS Aequorea victoria.
XX
PN US2004142433-A1.
XX
PD 22-JUL-2004.
XX
PF 10-OCT-2003; 2003US-00684134.
XX
PR 02-FEB-2001; 2001US-0266386P.
PR 14-FEB-2001; 2001US-0268785P.
PR 01-FEB-2002; 2002US-00066390.
PR 08-AUG-2002; 2002US-0402342P.
PR 21-AUG-2002; 2002US-00226372.
PR 25-OCT-2002; 2002US-00280913.
PR 08-AUG-2003; 2003US-00637758.
XX
PA (PADG/) PADGETT H S.

Db 2042 CACATTAATTCGGCTCCTCCTCCGCTTCCAGTCGGGAACCTGCTGCGCAGCT 1983
QY 1681 GCATTAATGAATCGCCACCGCGGGAGAGCGGTTTGGCTATTTGGCGCTCTTCCG 1740
Db 1982 GCATTAATGAATCGGCGCAACCGCGGGAGAGCGGTTTGGCTATTTGGCGCTCTTCCG 1923
QY 1741 TTCTCTCGCTCACTGACTCGCTCGCTCGGTCGGTTCGGCTCGCGCGAGCGGTATCAGCTCA 1800
Db 1922 TTCTCTCGCTCACTGACTCGCTCGCTCGGTCGGTTCGGCTCGCGCGAGCGGTATCAGCTCA 1863
QY 1801 CTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATTAACGACGAGAAACATGTG 1860
Db 1862 CTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATTAACGACGAGAAACATGTG 1803
QY 1861 AGCAAAAGCCAGCAAAAGGCCAGGAACCGTAAAGGCCGGTTGCTGGCGTTTTCCTCA 1920
Db 1802 AGCAAAAGCCAGCAAAAGGCCAGGAACCGTAAAGGCCGGTTGCTGGCGTTTTCCTCA 1743
QY 1921 TAGGCTCCGCCCTCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGCGGAAA 1980
Db 1742 TAGGCTCCGCCCTCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGCGGAAA 1683
QY 1981 CCCGACAGACTATAAAGATACAGGCGTTTCCCTCTGGAAGCTCCCTGCGGCTCTCC 2040
Db 1682 CCCGACAGACTATAAAGATACAGGCGTTTCCCTCTGGAAGCTCCCTGCGGCTCTCC 1623
QY 2041 TGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCCTCTCGGGAAGCGTGGC 2100
Db 1622 TGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCCTCTCGGGAAGCGTGGC 1563
QY 2101 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGTTCCGCTCAAGCT 2160
Db 1562 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGTTCCGCTCAAGCT 1503
QY 2161 GGGCTGTGTGCAAGAACCCCGCTTACGCCGACCGCTGCGCTTATCCGTTAACTATCG 2220
Db 1502 GGGCTGTGTGCAAGAACCCCGCTTACGCCGACCGCTGCGCTTATCCGTTAACTATCG 1443
QY 2221 TCTTCAGTCCACCGGTAAAGCAGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAG 2280
Db 1442 TCTTCAGTCCACCGGTAAAGCAGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAG 1383
QY 2281 GATTAGCAGAGCGAGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACTA 2340
Db 1382 GATTAGCAGAGCGAGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACTA 1323
QY 2341 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGGAAGCCAGTTACCTCGG 2400
Db 1322 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGGAAGCCAGTTACCTCGG 1263
QY 2401 AAAAGAGTTGGTAGCTCTTGATCCGGCAACAAACCCACCGCTGCTAGCGGTGGTTTTTT 2460
Db 1262 AAAAGAGTTGGTAGCTCTTGATCCGGCAACAAACCCACCGCTGCTAGCGGTGGTTTTTT 1203
QY 2461 TGTTTGAAGCAGCAGATTACCGCGAGAAAAAGGATCTCAAGAAGATCTCTTTGATCTTT 2520
Db 1202 TGTTTGAAGCAGCAGATTACCGCGAGAAAAAGGATCTCAAGAAGATCTCTTTGATCTTT 1143
QY 2521 TTCTACGGGCTGACGCTCAGTGGAAAGAAACCTCAGCTTAAAGGATTTTGGTCATGAG 2580
Db 1142 TTCTACGGGCTGACGCTCAGTGGAAAGAAACCTCAGCTTAAAGGATTTTGGTCATGAG 1083
QY 2581 ATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAATCAAGTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAATCAAGTTTAAATCAAT 1023
QY 2641 CTAAGATATATAGTAAACTTTGGTCTGACAGTTTACCAATGCTTAACTCAGTGGGCACC 2700
Db 1022 CTAAGATATATAGTAAACTTTGGTCTGACAGTTTACCAATGCTTAACTCAGTGGGCACC 963
QY 2701 TATCTCAGCGATCTGCTATTTGTTGTTTCATTCAGTGTGCTGACCTCCCGTCTGTAGAT 2760
|||||

Db 962 TATCTCAGCGATCTGCTATTTTGGTTTCATCCATAGTGTGCTGACCTCCCGTCTGTAGAT 903
QY 2761 AACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGTGCAATATGATACCCGAGACCC 2820
Db 902 AACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGTGCAATATGATACCCGAGACCC 843
QY 2821 AGCTCACCGGCTCAGATTTTATCAGCAATAAACAGCAGCCGGAAGGCCGAGCGCAG 2880
Db 842 AGCTCACCGGCTCAGATTTTATCAGCAATAAACAGCAGCCGGAAGGCCGAGCGCAG 783
QY 2881 AAGTGGTCTCGCAACTTTATCCGCCCTCATCCAGTCTATTAATTTGTTGCCGGGAAGCTAG 2940
Db 782 AAGTGGTCTCGCAACTTTATCCGCCCTCATCCAGTCTATTAATTTGTTGCCGGGAAGCTAG 723
QY 2941 AGTAAGTAGTTTCGCCAGTTAATAGTTTGCAGAACGTTTGTGCCATTTGCTACAGGATCGT 3000
Db 722 AGTAAGTAGTTTCGCCAGTTAATAGTTTGGGCAACGTTTGTGCCATTTGCTACAGGATCGT 663
QY 3001 GGTGTACGCTCGTTCGTTTGGTATGCTTTCATTTACGCTCCGTTCCCAACGATCAAGGCG 3060
Db 662 GGTGTACGCTCGTTCGTTTGGTATGCTTTCATTTACGCTCCGTTCCCAACGATCAAGGCG 603
QY 3061 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGT 3120
Db 602 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGT 543
QY 3121 TGTCAAGAAGTAAGTTGGCGCAGTGTATCACTCATGTTATGCGCAGCACTGCATAATTC 3180
Db 542 TGTCAAGAAGTAAGTTGGCGCAGTGTATCACTCATGTTATGCGCAGCACTGCATAATTC 483
QY 3181 TCTTACTGTCATGCGCATCCGTAAGATGCTTTTCTGTGATCGTGTAGTACTCAACCAAGTC 3240
Db 482 TCTTACTGTCATGCGCATCCGTAAGATGCTTTTCTGTGATCGTGTAGTACTCAACCAAGTC 423
QY 3241 ATTCTGAGAAATAGTGTATGCGCGCAGCCAGTGTCTTTGCCCGCGCTCAATACGGGATTA 3300
Db 422 ATTCTGAGAAATAGTGTATGCGCGCAGCCAGTGTCTTTGCCCGCGCTCAATACGGGATTA 363
QY 3301 TACCGCGCACATAGCAGAACTTTAAAAAGTGTCACTCATTTGGAAAACTTTCTCGGGCG 3360
Db 362 TACCGCGCACATAGCAGAACTTTAAAAAGTGTCACTCATTTGGAAAACTTTCTCGGGCG 303
QY 3361 AAAACTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTCGTGCACC 3420
Db 302 AAAACTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTCGTGCACC 243
QY 3421 CAACTGATCTTCAGCATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAAACAGGAG 3480
Db 242 CAACTGATCTTCAGCATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAAACAGGAG 183
QY 3481 GCAAAATCCGCAAAAAAGGGAATAAGGCGCAGACGGAAATGTTGAATACTCATACTCTT 3540
Db 182 GCAAAATCCGCAAAAAAGGGAATAAGGCGCAGACGGAAATGTTGAATACTCATACTCTT 123
QY 3541 CCTTTTCAATATTTATGAAGCATTTTATCAGGGTTATTTGCTCATGACGGGATACATATT 3600
Db 122 CCTTTTCAATATTTATGAAGCATTTTATCAGGGTTATTTGCTCATGACGGGATACATATT 63
QY 3601 TGAATGATTTAGAAAAATAAATAAGGGGTTCCGCGCACATTTCCCGCAAAAGTGCC 3660
Db 62 TGAATGATTTAGAAAAATAAATAAGGGGTTCCGCGCACATTTCCCGCAAAAGTGCC 3
QY 3661 AC 3662
|||
Db 2 AC 1
RESULT 10
ADSL1717/c
ID ADSL1717 standard; DNA; 3637 BP.
XX
AC ADSL1717;
XX

02-DEC-2004 (first entry)
pBSC3BFP plasmid DNA.
Polymerase; cyclic; circular; ds.
Unidentified.
US2004180352-A1.
16-SEP-2004.
08-AUG-2003; 2003US-00637758.
08-AUG-2002; 2002US-0402342P.
(LARG-) LARGE SCALE BIOLOGY CORP.
Padgett HS, Lindbo JA, Fitzmaurice WP, Vaewhongs AA;
WPI; 2004-667658/65.
In vitro method of making sequence variants from heteroduplex
polynucleotide, involves combining polynucleotide with agent having
polymerase and strand cleavage activities for sufficient time for
percentage of complementarity to increase.
Example 14; SEQ ID NO 17; 81pp; English.
The invention relates to an in vitro method of making sequence variants
from heteroduplex polynucleotide which involves combining polynucleotide
with an agent having polymerase and strand cleavage activities for
sufficient time for the percentage of complementarity to be increased
within the heteroduplex. The method is useful for performing in vitro
method of making sequence variants from one or more heteroduplex
polynucleotide. It is useful in increasing diversity in a population of
sequences and obtaining a polynucleotide encoding a desired functional
property. The present sequence is a pBSC3BFP plasmid DNA. This sequence
is used to illustrate the method of the invention.
Sequence 3637 BP; 942 A; 892 C; 870 G; 933 T; 0 U; 0 Other;
Query Match 72.8%; Score 2664.8; DB 13; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;
1 CTAATTTGTAAGCGTTAATTTTGTAAATTCGGTAAATTTTGTAAATTCAGCTC 60
3637 CTAATTTGTAAGCGTTAATTTTGTAAATTCGGTAAATTTTGTAAATTCAGCTC 3578
61 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 120
3577 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 3518
121 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAGAAACGTGGACTC 180
3517 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAGAAACGTGGACTC 3458
181 CAAGCTCAAGAGCGGAAACCGCTCTATCAGGCGATGGCCCACTACGTGAACCATCACC 240
3457 CAAGCTCAAGAGCGGAAACCGCTCTATCAGGCGATGGCCCACTACGTGAACCATCACC 3398
241 CTAATCAAGTTTGTGGGTTCGAGTGCCTTAAGCACTAAATCGGAACCTTAAGGGAG 300
3397 CTAATCAAGTTTGTGGGTTCGAGTGCCTTAAGCACTAAATCGGAACCTTAAGGGAG 3338
301 CCCCCGATTTAGAGTTGACGGGAAACCGCGGAAACGTGGCGAGAAAGGAAGGAGAA 360
3337 CCCCCGATTTAGAGTTGACGGGAAACCGCGGAAACGTGGCGAGAAAGGAAGGAGAA 3278
361 AGCGAAGGACGGGCGCTAGGCGGTGGCAAGTGTAGCGGTACCGTGGCGGTAAACCAC 420
3277 AGCGAAGGACGGGCGCTAGGCGGTGGCAAGTGTAGCGGTACCGTGGCGGTAAACCAC 3218

QY 421 CACACCCGCGCGCTTAATTCGCGCTACAGGGCGGTCCCATTCGCGCATTCAGGCTGCG 480
DB 3217 CACACCCGCGCGCTTAATTCGCGCTACAGGGCGGTCCCATTCGCGCATTCAGGCTGCG 3158
QY 481 CAACTGTTGGGAGGGCGGATCGGTGGGCGCTCTTCGCTATTACGCCAGCTGGCGAAGG 540
DB 3157 CAACTGTTGGGAGGGCGGATCGGTGGGCGCTCTTCGCTATTACGCCAGCTGGCGAAGG 3098
QY 541 GGGATGTCCTGCAAGCGGATTAAGTTGGGTAAAGCCAGGCGTTTCCAGTCAGACGTTG 600
DB 3097 GGGATGTCCTGCAAGCGGATTAAGTTGGGTAAAGCCAGGCGTTTCCAGTCAGACGTTG 3038
QY 601 TAAACGACGGCCAGTGAGCGCGCTCGTTCAATTCAGGTTTGTGAACCCCGTGGAGACGG 660
DB 3037 TAAACGACGGCCAGTGAGCGCGCTCACTATAGGGCGAATTTGGAGCTCCA 2978
QY 661 GCAGACTCGCGGTGCAAAATGTTTACAGCGGTGATGGAGCAGATGAAGATGTCGACAC 720
DB 2977 CGCGGTGGCG-----CGCTCTAGAACTAGTGGATCCCC--CGG 2939
QY 721 GCTGCAGAACACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTCAGTTAA 780
DB 2938 GCTGCAGGAATCTTATTGTTGATAGTTTCATCGCATGCTGTAATCCCGACGAGCTTA 2879
QY 781 AGATAATCATCGGTAAATTTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAAT 840
DB 2878 CAAACTCAAGAGAGGACCATGTGTCAGCGCTTTTCGTTGGGATCTTTCGAAAGGCGAGATT 2819
QY 841 TGGGTACCGGGCCCCCTCGAGTCGCGGTATCGAATAGCTTGATTCGAATTCCTGC 900
DB 2818 GTGTCGACAGGTAATGTTGTTCTGTTAAAGGACAGGGCCATCGCAATTTGGAGTATTTT 2759
QY 901 AGCCCGGGGATCCACTAGTTCTAGAGCGCGCCACCGCGTGGAGCTCCAGCTTTGT 960
DB 2758 GTTGATAATGGTCTGCTAGTTGNAACCGATCCATCTTCAATGTTGTGGCAATTTGAAT 2699
QY 961 TCCCTTTAGTAGGGTTAATAGATCCCATGCGTCAATTTTACGACAGACTATCTTCTAG 1020
DB 2698 TAGCTTTGATTCATCTTTTGTCTGCGCGTATGATGATATACATGTTGTGAG--TTATAG 2640
QY 1021 GGTATCTAGCTGCATCAGGATCATATCGTGGGTCTTTTTCGGCTCAGTCACTGCC 1080
DB 2639 TTGTACTCAGTTTGTGTCGAGAAATGTTTCCATCTTCTTAAATAACAATCTTTTAAAC 2580
QY 1081 CAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGACCGGTGCTTTTCCCGGAGGTG 1140
DB 2579 TCGATACGATTAAAGGGGTATCACCTTCNAATTCAGCTTCAGCAGCGCTCTTGTAGTTC 2520
QY 1141 AAGCGCATGGAAAGAGTTTGGCGAGGATGACTGCTGCTGCTGATGACGTTGAGCGAAGAAC 1200
DB 2519 CGGTCACTTTGAAAGATATAGTGCCTTCTGTACATAACCTTCGGGCATGGCACTCTTG 2460
QY 1201 GCAGCTTTACCATGATGATTCGGGAAGGTGTGGCCATGACGCGCTTTACGGTGAAGTGT 1260
DB 2459 AAAAACTCATCGCGTTTCATATGATCCGGATAACGGGAAAGCAATGAAACCATTAAGAG 2400
QY 1261 TCGTTCAAGGCACCTCGGATACAGTTCTGCGCGCTTTTCCGGACACAGTTCGGATGG 1320
DB 2399 AAGTAGTACAAGTGTGGCCATGGAAAGGTAGTTTCCAGTAGTGCAATAAATTTA 2340
QY 1321 TCAGCCCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACACCGGGAATCGCGTGGCG 1380
DB 2339 AGGGTAAGCTTTCCGTATGTAGCATCACCTTCACCTCTCCACTGACAGAAAAATTTGTGC 2280
QY 1381 GTGTGAGATTATGACAGGGTGGCGCTGGGATATTACGTACGGAGGACGGGTATC 1440
DB 2279 CCATTAACATCACTAATTTAAATTTGGGACAACTCCAGTGAAAGTTCTTCT 2220
QY 1441 CTGGCTGGGATGCGCGAAGAAATGGACATGGATACCCGTGAGTTTACCGGCGGCGCGCTT 1500
DB 2219 C-----CTTTACTCATCGGTACCCAGCTTTTGTTCCTTTTAGTAGGGGTAAATTCGCGCTT 2163

QY 1501 GCGTAAATCATGGTCATAGCTGTTCTCTGTGTGAATTTGTTATCCGCTCACAAATCCACA 1560
DB 2162 GCGTAAATCATGGTCATAGCTGTTCTCTGTGTGAATTTGTTATCCGCTCACAAATCCACA 2103
QY 1561 CAACATACGCGGAGCATAAAGTGTAAAGCTGGGGTGCCTAATAGTAGCTAACT 1620
DB 2102 CAACATACGCGGAGCATAAAGTGTAAAGCTGGGGTGCCTAATAGTAGCTAACT 2043
QY 1621 CACATTAATTTGGTGTGGCTCACTGCGCGCTTCCAGTCCGGAACCTCTGTCGCCAGCT 1680
DB 2042 CACATTAATTTGGTGTGGCTCACTGCGCGCTTCCAGTCCGGAACCTCTGTCGCCAGCT 1983
QY 1681 GCATTAATGAATCGGCACACGCGCGGGAGAGCGGTTTGCCTATTTGGCGCTCTTCCGC 1740
DB 1982 GCATTAATGAATCGGCACACGCGCGGGAGAGCGGTTTGCCTATTTGGCGCTCTTCCGC 1923
QY 1741 TTTCTCGCTCACTGACTCGCTGGGCTCGGTGTTCCAGTCCGGAACCTCTGTCGCCAGCTCA 1800
DB 1922 TTTCTCGCTCACTGACTCGCTGGGCTCGGTGTTCCAGTCCGGAACCTCTGTCGCCAGCTCA 1863
QY 1801 CTCAAGGCGGTAAATACGGTATTACACAGAATCAGGGGATAACGACAGGAAGAAACATGTG 1860
DB 1862 CTCAAAGGCGGTAAATACGGTATTACACAGAATCAGGGGATAACGACAGGAAGAAACATGTG 1803
QY 1861 AGCAAAAGCCAGCAAAAGGCCAGGAACCGTTAAAGGCGGCTTCTGCGCGTTTTCCTCA 1920
DB 1802 AGCAAAAGCCAGCAAAAGGCCAGGAACCGTTAAAGGCGGCTTCTGCGCGTTTTCCTCA 1743
QY 1921 TAGGCTCGCGCCCTTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGGAAA 1980
DB 1742 TAGGCTCGCGCCCTTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGGAAA 1683
QY 1981 CCCGACAGACTATAAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTCGTCGCTCTCC 2040
DB 1682 CCCGACAGACTATAAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTCGTCGCTCTCC 1623
QY 2041 TGTTCCGACCTGCGCTTACCGBATACCTGTCCGCTTTTCCCTTCGGGAAGCGTGCG 2100
DB 1622 TGTTCCGACCTGCGCTTACCGBATACCTGTCCGCTTTTCCCTTCGGGAAGCGTGCG 1563
QY 2101 GCTTTTCTCATAGCTCACGCTGTAGTATCTCAGTTTCGGTGTAGGTTCGTCCTCAAGCT 2160
DB 1562 GCTTTTCTCATAGCTCACGCTGTAGTATCTCAGTTTCGGTGTAGGTTCGTCCTCAAGCT 1503
QY 2161 GGGCTGTGTGACGAACCCCGCTTACAGCCGACCGCTGCGCTTATCCGGTAACTATCG 2220
DB 1502 GGGCTGTGTGACGAACCCCGCTTACAGCCGACCGCTGCGCTTATCCGGTAACTATCG 1443
QY 2221 TCTTGAGTCCAAACCCGGTAAAGACACGACTTATGCGCACTGGCAGCAGCACTGGTAAACAG 2280
DB 1442 TCTTGAGTCCAAACCCGGTAAAGACACGACTTATGCGCACTGGCAGCAGCACTGGTAAACAG 1383
QY 2281 GATTAGCAGCGAGGTATGTAGGCGGTGCTACAGATTCTTGAAGTGGTGGCTTAACTA 2340
DB 1382 GATTAGCAGCGAGGTATGTAGGCGGTGCTACAGATTCTTGAAGTGGTGGCTTAACTA 1323
QY 2341 CGGCTACACTAGAAGCAGAGTATTTGGTATCTGCGCTCTGCTGGAAGCCAGTTTACCTTCGG 2400
DB 1322 CGGCTACACTAGAAGCAGAGTATTTGGTATCTGCGCTCTGCTGGAAGCCAGTTTACCTTCGG 1263
QY 2401 AAAAAAGTTGGTAGCTCTTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGGTTTTTT 2460
DB 1262 AAAAAAGTTGGTAGCTCTTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGGTTTTTT 1203
QY 2461 TGTTTGCAAGCAGCAGATTACCGGCAGAAAAAAGGATCTCAAGAAGATCTTTGATCTT 2520
DB 1202 TGTTTGCAAGCAGCAGATTACCGGCAGAAAAAAGGATCTCAAGAAGATCTTTGATCTT 1143
QY 2521 TTTCTACGGGCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGATTTTGGTCATGAG 2580
DB 1142 TTTCTACGGGCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGATTTTGGTCATGAG 1083
QY 2581 ATTATCAAAAAGGATCTTCACTACCTAGATCCTTTTAAATTTAAAAATGAAATCAAT 2640

DB 1082 ATTATCAAAAAGGATCTTCACTACCTAGATCCTTTTAAATTTAAAAATGAAATCAAT 1023
QY 2641 CTAAAGTATATATAGTAAACTTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGAGGCACC 2700
DB 1022 CTAAAGTATATATAGTAAACTTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGAGGCACC 963
QY 2701 TATCTCAGCGATCTCTGTATTTTGGTTCATCCATAGTTGCTGACTCCCCGCTCGTGTAGAT 2760
DB 962 TATCTCAGCGATCTCTGTATTTTGGTTCATCCATAGTTGCTGACTCCCCGCTCGTGTAGAT 903
QY 2761 AACTACGATACCGGAGGGCTTACCATCTG6CCCCAGTGTGCAATGATACCGGAGACCC 2820
DB 902 AACTACGATACCGGAGGGCTTACCATCTG6CCCCAGTGTGCAATGATACCGGAGACCC 843
QY 2821 AGGCTCACCGGCTCCAGATTTTATCAGCAATAAACAGCCAGCGGAAGGGCGAGCGCAG 2880
DB 842 AGGCTCACCGGCTCCAGATTTTATCAGCAATAAACAGCCAGCGGAAGGGCGAGCGCAG 783
QY 2881 AAGTGGTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGAAGCTAG 2940
DB 782 AAGTGGTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGAAGCTAG 723
QY 2941 AGTAAAGTATTGCGCAGTTAATAGTTTGGCAACGTTGTGTCATTTGTCAGGCACTGT 3000
DB 722 AGTAAAGTATTGCGCAGTTAATAGTTTGGCAACGTTGTGTCATTTGTCAGGCACTGT 663
QY 3001 GGTGTACGCTCGTCTGTTGGTATGCTTCAATTCAGCTCCGTTCCCAACGATCAAGGG 3060
DB 662 GGTGTACGCTCGTCTGTTGGTATGCTTCAATTCAGCTCCGTTCCCAACGATCAAGGG 603
QY 3061 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGCTCCTTCGGTCTCCGATCGT 3120
DB 602 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGCTCCTTCGGTCTCCGATCGT 543
QY 3121 TGTCAAGATGAAGTTGGCGCAGTGTATCACTCATGTTTATGGCAGCAGCTGCATAATTC 3180
DB 542 TGTCAAGATGAAGTTGGCGCAGTGTATCACTCATGTTTATGGCAGCAGCTGCATAATTC 483
QY 3181 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTAAGTGTGTAAGTCTCAACCAAGTC 3240
DB 482 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTAAGTGTGTAAGTCTCAACCAAGTC 423
QY 3241 ATTCTGAGAATAGTGTATGCGCGACCGAGTTGCTCTTTGCCGCGGTCAATACGGGATAA 3300
DB 422 ATTCTGAGAATAGTGTATGCGCGACCGAGTTGCTCTTTGCCGCGGTCAATACGGGATAA 363
QY 3301 TACCGCGCACATGACGAACTTTTAAAGTGTCTCATCATTTGGAAAAACGTTCTTCGGGGCG 3360
DB 362 TACCGCGCACATGACGAACTTTTAAAGTGTCTCATCATTTGGAAAAACGTTCTTCGGGGCG 303
QY 3361 AAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATTAACCCACCTCGTGACCC 3420
DB 302 AAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATTAACCCACCTCGTGACCC 243
QY 3421 CAACTGATCTTCAGCATCTTTTACTTTTCCAGCGCTTTCTGGGTGAGCAAAAACAGGAAG 3480
DB 242 CAACTGATCTTCAGCATCTTTTACTTTTCCAGCGCTTTCTGGGTGAGCAAAAACAGGAAG 183
QY 3481 GCAAAATCCGCAAAAAGGGAATAAGGGCGACACGGAATAATGTTGAATACTCATCTCTT 3540
DB 182 GCAAAATCCGCAAAAAGGGAATAAGGGCGACACGGAATAATGTTGAATACTCATCTCTT 123
QY 3541 CCTTTTCAATATTTATGAAGCAATTTATCAGGGTATTGCTCATGAGCGGATACATATT 3600
DB 122 CCTTTTCAATATTTATGAAGCAATTTATCAGGGTATTGCTCATGAGCGGATACATATT 63
QY 3601 TGAATGATTTTGAAGCAATTTATCAGGGTATTGCTCATGAGCGGATACATATT 3660
DB 62 TGAATGATTTTGAAGCAATTTATCAGGGTATTGCTCATGAGCGGATACATATT 3
QY 3661 AC 3662

```
Db          2 AC 1
RESULT 11
ABX14570/c
ID ABX14570 standard; DNA; 3637 BP.
XX
XX
AC ABX14570;
XX
XX 14-MAR-2003 (first entry)
XX
XX Plasmid pBSWTFPP DNA.
XX
XX Circular; cyclic; heteroduplex; T4 DNA polymerase; T4 DNA ligase; CEL I;
KW population diversity; exonuclease; complementarity; ds.
KW
XX Synthetic.
OS
XX
XX WO200279468-A2.
XX
XX 10-OCT-2002.
XX
XX 01-FEB-2002; 2002WO-US003055.
XX
XX 02-FEB-2001; 2001US-0266386P.
PR
XX 14-FEB-2001; 2001US-0268785P.
XX
XX (LARG-) LARGE SCALE BIOLOGY CORP.
PA
XX Padgett HS, Fitzmaurice WP, Lindo JA;
PI WPI; 2003-092898/08.
DR
XX
XX Increasing complementarity by making sequence variants from heteroduplex
PT polynucleotides, useful for producing improved gene products from
PT randomly mutated genes.
XX
XX Example 5; Page 98-99; 105pp; English.
XX
XX The invention relates to an in vitro method of making sequence variants
CC from at least one heteroduplex polynucleotide that has at least two-non
CC complementary nucleotide base pairs. The method comprises combining a
CC heteroduplex polynucleotide with CEL I, T4 DNA polymerase and T4 DNA
CC ligase, or an agent or agents with exonuclease activity and allowing
CC sufficient time for the percentage of complementarity to increase, where
CC one or more variants are made. The method can also be used to increase
CC diversity in a population of sequences. The method is useful for
CC producing improved gene products from randomly mutated genes or from
CC expression from a suitable plant, animal, fungal, yeast or bacterial
CC expression vector, and provides a high-efficiency recovery of
CC recombinants for generating novel polynucleotides from parental templates
CC that are more diverse and with a lower percentage of sequence identity.
CC This sequence represents plasmid DNA, used in heteroduplex substrate
CC preparation
XX
XX Sequence 3637 BP; 959 A; 879 C; 868 G; 931 T; 0 U; 0 Other;
SQ
Query Match 72.5%; Score 2656.4; DB 8; Length 3637;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;
QY 1 CTAATTTGTAAGCGTTAAATTTTGTGTTAAATTCGCGTTAAATTTTGTGTTAAATTCAGCTC 60
DB 3637 CTAATTTGTAAGCGTTAAATTTTGTGTTAAATTCGCGTTAAATTTTGTGTTAAATTCAGCTC 3578
QY 61 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 120
DB 3577 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 3518
QY 121 GATAGGGTTGAGTGTGTTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACTGGACTC 180
DB 3517 GATAGGGTTGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACTGGACTC 3458
```


QY 3479 AGCAAAATGCCCAAAAAAGGAATAAGGCGACACGGAATGTTGAATACTCATACTC 3538
DB 184 AGCAAAATGCCCAAAAAAGGAATAAGGCGACACGGAATGTTGAATACTCATACTC 125
QY 3539 TTCTCTTTTCAATATTATTGAAGCATTTATCAGGTTTATGTTCTCATGAGCGATACATA 3598
DB 124 TTCTCTTTTCAATATTATTGAAGCATTTATCAGGTTTATGTTCTCATGAGCGATACATA 65
QY 3599 TTGGAATGTATTTAGAAAAATAAAATAGGGGTTCCGCGCACATTTCCCGGAAAAAGTG 3658
DB 64 TTGGAATGTATTTAGAAAAATAAAATAGGGGTTCCGCGCACATTTCCCGGAAAAAGTG 5
QY 3659 CCAC 3662
DB 4 CCAC 1

RESULT 12
ADM68462/C
ID ADM68462 standard; DNA; 3637 BP.
XX
AC ADM68462;
XX
DT 03-JUN-2004 (first entry)
DE Wild type green fluorescent protein, GFP, gene construct.
XX
KW ds; mismatch endonuclease; endonuclease; gene shuffling technology;
KW single nucleotide polymorphism; cancer susceptibility;
KW sequence variation redistribution; GFP; green fluorescent protein; gene.
XX
OS Aequorea victoria.
OS Synthetic.
XX
PN US2003157682-A1.
XX
PD 21-AUG-2003.
XX
PF 31-JAN-2003; 2003US-00356708.
PR 01-FEB-2002; 2002US-0353722P.
PR 14-MAR-2002; 2002US-00098155.
PR 01-AUG-2002; 2002US-00211079.
XX
PA (PADG/) PADGETT H S.
PA (VAEW/) VAETHONGS A A.
PA (VOJD/) VOJDANI F S.
PA (SMIT/) SMITH M L.
PA (LIND/) LINDBO J A.
PA (FITZ/) FITZMAURICE W P.
XX
PI Padgett HS, Vaethongs AA, Vojdani FS, Smith ML, Lindbo JA;
PI Fitzmaurice WP;
XX
DR WPI; 2003-766176/72.
XX
PT Making a mismatch endonuclease, useful in gene shuffling and in detection
PT of single nucleotide polymorphisms, comprises transfecting a host with a
PT recombinant viral vector including a polynucleotide encoding a mismatch
PT endonuclease.
XX
PS Example 5; SEQ ID NO 31; 79pp; English.
XX
CC The invention relates to a method of making a mismatch endonuclease
CC enzyme comprising transfecting a host plant, animal, yeast, fungus or
CC bacterium with a recombinant viral vector that encodes a polynucleotide
CC sequence for a mismatch endonuclease, growing the host so that the
CC polynucleotide is expressed, and extracting the mismatch endonuclease
CC enzyme from the host. The method is useful for making mismatch
CC endonuclease enzymes, for obtaining peptides and polynucleotides with
CC desired functional properties and for detecting mutations. The mismatch
CC endonuclease enzymes are useful in gene shuffling technology for

CC developing new genes, in detecting single nucleotide polymorphisms for
CC e.g. detecting evidence of cancer susceptibility, or in redistributing
CC sequence variations between non-identical polynucleotide sequences. The
CC present sequence represents the wild type green fluorescent protein, GFP,
CC gene construct.
XX Sequence 3637 BP; 959 A; 879 C; 868 G; 931 T; 0 U; 0 Other;
SQ

Query Match 72.5%; Score 2656.4; DB 11; Length 3637;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;

QY 1 CTAAATTTAAGCGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTC 60
DB 3637 CTAAATTTAAGCGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTC 3578
QY 61 ATTTTAAACCAATAGCGCAAAATCGCAAAATCCCTTATAATCAAAAGNATAGACCGA 120
DB 3577 ATTTTAAACCAATAGCGCGCAAAATCGCAAAATCCCTTATAATCAAAAGNATAGACCGA 3518
QY 121 GATAGGTTGAGTTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTGACTC 180
DB 3517 GATAGGTTGAGTTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTGACTC 3458
QY 181 CAAGTCAAAAGGCGGCAAAACCGTCTATCAGGGCGATGCGCCCACTACGTGAACCATCAC 240
DB 3457 CAAGTCAAAAGGCGGCAAAACCGTCTATCAGGGCGATGCGCCCACTACGTGAACCATCAC 3398
QY 241 CTAAATCAAGTTTTTGGGTCGAGGTCGCGTAAAGCACTAAATCGGAACCTTAAGGGAG 300
DB 3397 CTAAATCAAGTTTTTGGGTCGAGGTCGCGTAAAGCACTAAATCGGAACCTTAAGGGAG 3338
QY 301 CCCCCGATTTAGAGCTTGAACGGGAAAGCCGCGAAACGTCGCGAGAGAAAGGAAGAA 360
DB 3337 CCCCCGATTTAGAGCTTGAACGGGAAAGCCGCGAAACGTCGCGAGAGAAAGGAAGAA 3278
QY 361 AGCGAAAGAGCGGCGCTAGGCGCTGCAAGTGTAGCGGTACACCTGGCGGTAAACAC 420
DB 3277 AGCGAAAGAGCGGCGCTAGGCGCTGCAAGTGTAGCGGTACACCTGGCGGTAAACAC 3218
QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGGCGCTCCCATTTGCCCATTTCAAGGTGCG 480
DB 3217 CACACCGCGCGCTTAATGCGCGCTACAGGGCGCTCCCATTTGCCCATTTCAAGGTGCG 3158
QY 481 CAACTGTTGGGAAGGCGCATCGGTGCGGCTCTTTCGCTATTACGCCAGTGGCGAAAGG 540
DB 3157 CAACTGTTGGGAAGGCGCATCGGTGCGGCTCTTTCGCTATTACGCCAGTGGCGAAAGG 3098
QY 541 GGGATGTGCTCAAGGCGATTAAGTTGGGTAAAGCGCAGGTTTCCAGTCAGCAGCTTG 600
DB 3097 GGGATGTGCTCAAGGCGATTAAGTTGGGTAAAGCGCAGGTTTCCAGTCAGCAGCTTG 3038
QY 601 TAAACGACGCGCAGTCAGCGCGCTTCGTTCAATTCACGTTTTTGAACCGTGGAGGCGG 660
DB 3037 TAAACGACGCGCAGTCAGCGCGCTTCGTTCAATTCACGTTTTTGAACCGTGGAGTCCA 2978
QY 661 GCAGACTCCGCTGCAAAATGTTTACAGCGTGTAGGAGCAGATCAAGATGCTCGACAC 720
DB 2977 CCGCGGTGCGCGCGCTCTAGAACTAGTGTGATCCCGCGGCTGCAGGAATCTTATTGT 2918
QY 721 GTGCGAAGACACGCGAGCTAGATAACCTTAGAAAGATAATCATATTGTGACGTACGTAA 780
DB 2917 ATAGTTTATCCATGTCATGTAAATCCAGCAGCTGTAACTCAAGAAAGGACCATGT 2858
QY 781 AGATAATC--ATGCGTAAATTTGACGCGATGGGATCTGTAATACGACTCACTATAGGCGA 838
DB 2857 GGTCTCTCTTTTTCGTTGGGATCTTTGCGAAAGGCGAGATTGTGTGGACAGGTAAATGTGT 2798
QY 839 ATTGGGTATCCGGGCCCCCTCCAGGTCCAGGTATCGATAGCTTGATATCGAAATCTCT 898
DB 2797 CTGGTAAAGGACAGGCGCCATCGCCAAATGGAGTATTTTGTGATTAATGGTCTGCTAGT 2738
QY 899 GCAGCCCGGGGATCCACTAGTTCTTAGAGCGGCGCCACCGCGGTGGAGCTCCAGCTTTT 958

Db 2737 GAAAGCTTCCATCTTCAATGTTGTGCTAAATTTTGAAGTTAACTTTTGATTCATCTTTT 2678
QY 959 GTTCCCTTTAGTGGGTTAAATTAGATCCATGCGTCAATTTTACGACACTATCTTTCT 1018
Db 2677 GTTGTGCTGCCAAGATGATACATTTGTGGA-GTTATAGTTGTTATTCCAAATTTGTGCA 2619
QY 1019 AGGGTTAAATCTAGCTGCATCAGGATCATATCGTGGGTCTTTTTCGCGCTCAGTCATCG 1078
Db 2618 AGAATGTTTCCATCTTCTTAAATCAATACCTTTTAACTCGAATC-----TATTAAANA 2564
QY 1079 CCCAAGCTGGCGCTATCTGGGCATCGGGAGGAAGAACCCGTGCTTTTCCGCGAGGT 1138
Db 2563 GGGTATCACCTTCAAATTTGACTTC-----AGCACGTGCTTGTAGTCCCGTC 2515
QY 1139 TGAAGCGGCATGGAAGAGTTTGGCGAGATGACTGCTGCTGCATTTGAGCGGAA 1198
Db 2514 ATCTTTGAAAAATATAGTTTCTTCTGTATACAACTTCGGGCATGGCACTCTTGAANA 2455
QY 1199 AGCGACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCTTTTAAACGGTGAAC 1258
Db 2454 GTCATGCCGTTTCATATGATCTGGGTACTTGAAGCAATTCGAACCATAGAGAAAGT 2395
QY 1259 GTTCGTTCAAGCCACTTGGGATACCAGTTTCGTGCGGCTTTTCCGGACACAGTTCCGGAT 1318
Db 2394 AGTGACAAGTGTGGCCATGGAACAGGTAGTTTTCAGTAGTGCAATAAAATTTAAGGT 2335
QY 1319 GGTACGCCAAGCGATCAGCAACCCGNAACAATACCGCGACAGCCGGAACCTGCGTGC 1378
Db 2334 AAGTTTTC-----CGTATGTTGCATCACCTTCACCTCTCCACTGACAGAAAAATTTGTGC 2280
QY 1379 CGGTGTGAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTACAGCGAGCGGTA 1438
Db 2279 CCAATTAACATCACATCTAATTCACAAGAAATTTGGACAACCTCCAGTGAAGAGTTCTTCT 2220
QY 1439 TCTCGCTGGATGSCCGCAGAAATGGACATGGATACCCCGTAGTTACCCGCGGCGCGC 1498
Db 2219 CTTTTACTCATCGGTACCCAGCTTTTGTGTTTCCCTTTAGTGAGGGTTA-----ATTGCGGC 2165
QY 1499 TTGGGTAATCATGTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCAATTTCCA 1558
Db 2164 TTGGCGTAATCATGTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCAATTTCCA 2105
QY 1559 CACAACATACGAGCGGAGCAATAAGTGTAAAGCTTGGGTCCTTAATGAGTAGCTAA 1618
Db 2104 CACAACATACGAGCGGAGCAATAAGTGTAAAGCTTGGGTCCTTAATGAGTAGCTAA 2045
QY 1619 CTCACATTAATTCGTTGCGCTCACTGCGCGCTTTTCCAGTCGGGAAACCTGTCGTGCCAG 1678
Db 2044 CTCACATTAATTCGTTGCGCTCACTGCGCGCTTTTCCAGTCGGGAAACCTGTCGTGCCAG 1985
QY 1679 CTGCATTAATGAATCGGCCAACCGCGGGGAGAGGGGTTTCGTTATTTGGGCGCTCTTCC 1738
Db 1984 CTGCATTAATGAATCGGCCAACCGCGGGGAGAGGGGTTTCGTTATTTGGGCGCTCTTCC 1925
QY 1739 GCTTCTCGCTCACTGACTCGCTGCGCTCGGTGCTTTCGCTGCGCGAGCGGTATCAGCT 1798
Db 1924 GCTTCTCGCTCACTGACTCGCTGCGCTCGGTGCTTTCGCTGCGCGAGCGGTATCAGCT 1865
QY 1799 CACTCAAGGGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACCGAGGAAGAACAATG 1858
Db 1864 CACTCAAGGGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACCGAGGAAGAACAATG 1805
QY 1859 TGAGCAAAAGGCGCAGAAAGGCGAGAACCGTAAAGGCGCGGTTCGTCGCGCTTTTTC 1918
Db 1804 TGAGCAAAAGGCGCAGAAAGGCGAGAACCGTAAAGGCGCGGTTCGTCGCGCTTTTTC 1745
QY 1919 CATAGGCTCCGCCCTTCAGCAGCATCAAAAAATTCGACGCTCAAGTCAGAGGTGGCGA 1978
Db 1744 CATAGGCTCCGCCCTTCAGCAGCATCAAAAAATTCGACGCTCAAGTCAGAGGTGGCGA 1685
QY 1979 AACCCGACAGGACTATAAGATACAGCGGTTTCCCGCTTGGAAAGTTCCTCGTGGCTCT 2038

Db 1684 AACCCGACAGGACTATAAAGATACACAGCGTTTCCCGCTGGAAGTCCCTCGTGGCTCT 1625
QY 2039 CCTGTTCCGACCGCTCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGCGGAAGCGTG 2098
Db 1624 CTTGTTCCGACCGCTCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGCGGAAGCGTG 1565
QY 2099 GCGCTTCTCATAGCTCAAGCTGTAGGTATCTCAGTTCGGTGTAGGTGTTTCGCTCCAAG 2158
Db 1564 GCGCTTCTCATAGCTCAAGCTGTAGGTATCTCAGTTCGGTGTAGGTGTTTCGCTCCAAG 1505
QY 2159 CTGGGCTGTGTGCAAGAACCCCGCTTCAAGCCGACCGCTGCGCTTATTCGCGTAACAT 2218
Db 1504 CTGGGCTGTGTGCAAGAACCCCGCTTCAAGCCGACCGCTGCGCTTATTCGCGTAACAT 1445
QY 2219 GGTCTTGTAGTCAACCCCGTAAAGACACACTTATCGCACTGGCAGCAGCCTCTGTAAC 2278
Db 1444 GGTCTTGTAGTCAACCCCGTAAAGACACACTTATCGCACTGGCAGCAGCCTCTGTAAC 1385
QY 2279 AGGATTAGCAGAGCGAGGTATGTAGGCGTGTACAGAGTCTTGAAGTGTGGCTTAAC 2338
Db 1384 AGGATTAGCAGAGCGAGGTATGTAGGCGTGTACAGAGTCTTGAAGTGTGGCTTAAC 1325
QY 2339 TAGGCTACATAGAAAGACAGTATTTGGTATCTCGCTCTGCTGAGCCAGTACCTTTC 2398
Db 1324 TAGGCTACATAGAAAGACAGTATTTGGTATCTCGCTCTGCTGAGCCAGTACCTTTC 1265
QY 2399 GGAAGAGAGTTGTGTAGCTCTTGATCCGCAACAAACACCGCTGGTAGCGTGTGTTT 2458
Db 1264 GGAAGAGAGTTGTGTAGCTCTTGATCCGCAACAAACACCGCTGGTAGCGTGTGTTT 1205
QY 2459 TTTGTTTCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAAGATCCTTTGATC 2518
Db 1204 TTTGTTTCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAAGATCCTTTGATC 1145
QY 2519 TTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAACTCAGTGTAAAGGATTTTGTTCATG 2578
Db 1144 TTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAACTCAGTGTAAAGGATTTTGTTCATG 1085
QY 2579 AGATTATCAAAAGAGTCTTCACTAGATCTTTTAAATTTAAATTTAAATTTAAATCA 2638
Db 1084 AGATTATCAAAAGAGTCTTCACTAGATCTTTTAAATTTAAATTTAAATTTAAATCA 1025
QY 2639 ATCTAAAGTATATATGAGTAAACTTTGGTCTGACAGTTACCAATGTCTTAATCAGTGAGCA 2698
Db 1024 ATCTAAAGTATATATGAGTAAACTTTGGTCTGACAGTTACCAATGTCTTAATCAGTGAGCA 965
QY 2699 CCTATCTCAGCGATCTGCTATTTTCGTTTCATCCATAGTTGCTGACTCCCCGCTGTTAG 2758
Db 964 CCTATCTCAGCGATCTGCTATTTTCGTTTCATCCATAGTTGCTGACTCCCCGCTGTTAG 905
QY 2759 ATAACTACGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATATCCGCGAGAC 2818
Db 904 ATAACTACGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATATCCGCGAGAC 845
QY 2819 CCACGCTCAACGGCTCCAGATTTTATCAGCAATAAACCAGCCAGCCGGAAGCCCGAGCGC 2878
Db 844 CCACGCTCAACGGCTCCAGATTTTATCAGCAATAAACCAGCCAGCCGGAAGCCCGAGCGC 785
QY 2879 AGAAGTGTCTGTGCAACTTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAGCT 2938
Db 784 AGAAGTGTCTGTGCAACTTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAGCT 725
QY 2939 AGAGTAAGTGTGTCCGAGTTAAATAGTTTGGCAACCTTTGGCCATTGCTACAGGCATC 2998
Db 724 AGAGTAAGTGTGTCCGAGTTAAATAGTTTGGCAACCTTTGGCCATTGCTACAGGCATC 665
QY 2999 GTGGTGTCAAGCTCGTGTGTTGGTATGGCTTCATTCAAGTCCGCTTCCCAACGATCAAGG 3058
Db 664 GTGGTGTCAAGCTCGTGTGTTGGTATGGCTTCATTCAAGTCCGCTTCCCAACGATCAAGG 605
QY 3059 CGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGTCTCTCGGTCTCTCGGATC 3118
Db 604 CGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGTCTCTCGGTCTCTCGGATC 545

```
QY 3119 GTTGTGAGAAAGTAAAGTGGCGGAGTGTATATCACTCATGTTATGGCAGCACTGCATAAT 3178
D 544 GTTGTGAGAAAGTAAAGTGGCGGAGTGTATATCACTCATGTTATGGCAGCACTGCATAAT 485
QY 3179 TCTCTTACTGTCTATGTCATCCGTAAGATGCTTTTCTGTGACTGCTGAGTACTCAACCAAG 3238
D 484 TCTCTTACTGTCTATGTCATCCGTAAGATGCTTTTCTGTGACTGCTGAGTACTCAACCAAG 425
QY 3239 TCATTCTGAGAAATAGTGTATGCGGACCGAGTGTGCTTTGCGCGGCGTCAATACGGGAT 3298
D 424 TCATTCTGAGAAATAGTGTATGCGGACCGAGTGTGCTTTGCGCGGCGTCAATACGGGAT 365
QY 3299 AATACCGCGCCACATAGCAGAACTTTAAAGTGCTCATATTGGAAAAACGTTCTTCGGGG 3358
D 364 AATACCGCGCCACATAGCAGAACTTTAAAGTGCTCATATTGGAAAAACGTTCTTCGGGG 305
QY 3359 CGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTCGTGCA 3418
D 304 CGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTCGTGCA 245
QY 3419 CCCAACTGATCTTCAGCATCTTTTACTTTTCCACGAGGCTTTCTGGGTGAGCAAAACAGGA 3478
D 244 CCCAACTGATCTTCAGCATCTTTTACTTTTCCACGAGGCTTTCTGGGTGAGCAAAACAGGA 185
QY 3479 AGCAAAATGCCCCAAAGGAATAAGGCGACACGGAAATGTTGAATACCTCATCTC 3538
D 184 AGCAAAATGCCCCAAAGGAATAAGGCGACACGGAAATGTTGAATACCTCATCTC 125
QY 3539 TTCTCTTTTCAATATTATTAAGCATTTATCAGGCTTATGCTCATGAGCGGATACATA 3598
D 124 TTCTCTTTTCAATATTATTAAGCATTTATCAGGCTTATGCTCATGAGCGGATACATA 65
QY 3599 TTTGAATGTATTTAGAAAAATAACAAATAGGGGTTCCGCGCACATTTCCCGGAAAAAGTG 3658
D 64 TTTGAATGTATTTAGAAAAATAACAAATAGGGGTTCCGCGCACATTTCCCGGAAAAAGTG 5
QY 3659 CCAC 3662
D 4 CCAC 1

RESULT 13
ADP26615/C
ID ADP26615 standard; DNA; 3637 BP.
XX
AC ADP26615;
XX
XX Green fluorescent protein (GFP) plasmid DNA #1.
XX
DE 26-AUG-2004 (first entry)
XX
XX Sequence variation; heteroduplex; transcription; DNA integration;
XX ribozyme expression; gene; ds; green fluorescent protein; GFP.
XX
OS Aequorea victoria.
OS Synthetic.
XX
XX US2004110130-A1.
XX
XX 10-JUN-2004.
XX
XX 25-OCT-2002; 2002US-00280913.
XX
XX 02-FEB-2001; 2001US-0266386P.
XX
XX 14-FEB-2001; 2001US-0268785P.
XX
XX 01-FEB-2002; 2002US-00066390.
XX
XX 08-AUG-2002; 2002US-0402342P.
XX
XX (LARG-) LARGE SCALE BIOLOGY CORP.
XX
XX Padgett HS, Lindbo JA, Fitzmaurice WP;
XX
```

```
DR WPI: 2004-440326/41.
XX Redistributing sequence variations between non-identical polynucleotide
PT sequences, useful for generating improved polynucleotide having a desired
PT characteristic, comprises making a heteroduplex and introducing a nick.
XX
XX Example 5; SEQ ID NO 31; 75bp; English.
XX
CC The invention relates to an in vitro method of redistributing sequence
CC variations between non-identical polynucleotide sequences, comprising
CC making a heteroduplex polynucleotide from two non-identical
CC polynucleotides, introducing a nick in the second strand at or near a
CC base pair mismatch site, removing the mismatched base(s) from the
CC mismatch site where the nick occurred and using the first strand as a
CC template to replace the removed base(s) with bases that complement the
CC base(s) in the first strand. The invention also relates to an in vitro
CC method of making a population of sequence variants from a heteroduplex
CC polynucleotide sequence, obtaining a polynucleotide sequence encoding a
CC desired functional property and identifying a reassorted DNA molecule
CC encoding a protein with a desired functional property. The method is
CC useful for generating an improved polynucleotide sequence or a population
CC of improved polynucleotide sequences possessing at least one desired
CC phenotypic characteristic (e.g., promotes transcription of linked
CC polynucleotides), where such polynucleotides are useful for expression
CC from a plant, animal, fungal, yeast, or bacterial expression vector, for
CC integration to form a transgenic plant, animal or microorganism, and for
CC expression of a ribozyme. This sequence represents DNA used in the scope
CC of the invention.
XX
SQ Sequence 3637 BP; 959 A; 879 C; 868 G; 931 T; 0 U; 0 Other;
Query Match 72.5%; Score 2656.4; DB 12; Length 3637;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;
QY 1 CTAATTTGTAACGGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTC 60
D 3637 CTAATTTGTAACGGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTC 3578
QY 61 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 120
D 3577 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 3518
QY 121 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGACTC 180
D 3517 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGACTC 3458
QY 181 CAACGTCAAAGGGCGGAAACCGCTCTATCAGGGCGATGCGCCACTACGTGAACCATCACC 240
D 3457 CAACGTCAAAGGGCGGAAACCGCTCTATCAGGGCGATGCGCCACTACGTGAACCATCACC 3398
QY 241 CTAATCAAGTTTTTTTGGGGTCGAGGTGCGTAAAGCACTAAATCGGAACCCCTAAAGGGAG 300
D 3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCGTAAAGCACTAAATCGGAACCCCTAAAGGGAG 3338
QY 301 CCCCCGATTTAGAGCTTGAACGGGGAAAGCCGGCGAAACGTCGGCGAGAAAGGAGGAGAA 360
D 3337 CCCCCGATTTAGAGCTTGAACGGGGAAAGCCGGCGAAACGTCGGCGAGAAAGGAGGAGAA 3278
QY 361 AGGGAAGAGCGGGCGCTAGGGCGCTGACAGTGTAGCGGTACCGTCCGCTGCGGTACCCAC 420
D 3277 AGGGAAGAGCGGGCGCTAGGGCGCTGACAGTGTAGCGGTACCGTCCGCTGCGGTACCCAC 3218
QY 421 CACACCGCGCGCTTAAATGCGCGCTACAGGGCGGTCCCATTCGCCATTTCAGGCTGCG 480
D 3217 CACACCGCGCGCTTAAATGCGCGCTACAGGGCGGTCCCATTCGCCATTTCAGGCTGCG 3158
QY 481 CAACTGTTGGGAAGGGCGATCGGTGCGGCGCTTTCGCTATTACGCCAGCTGCGGAAGG 540
D 3157 CAACTGTTGGGAAGGGCGATCGGTGCGGCGCTTTCGCTATTACGCCAGCTGCGGAAGG 3098
QY 541 GGGATGTGTCGAAGCGGATTAAGTTGGGTAAAGCCAGCGGTTTTTCCAGTCAGAGCTTG 600
D
```

Db	3097	GGGATGTGCTGCAAGGCGAATTAAGTTGGGTAAAGCCAGGCGTTTCCAGTCACGAGTTG	3038
Qy	601	TAAACGAGCGGCGAGTACGCGCGCTCGTTCAATCAAGTTTGAACCCGTCGAGACGG	660
Db	3037	TAAACGAGCGGCGAGTACGCGCGCTCGTTCAATCAAGTTTGAACCCGTCGAGACGG	2978
Qy	661	GCAGACTCCGCGTCAAAATGTGTTTACAGCGTGTAGAGCAGATGAAGATGCTCGACAC	720
Db	2977	CGCGGTGGCGCGCTCTAGAACTAGTGATCCCGCGGCTGCAGGAATCTTATTGT	2918
Qy	721	GCTGCAGAACACGCGAGCTAGATAACCCCTAGAAAAGATAATCATATGTGACGTACGTTAA	780
Db	2917	ATAGTTTCATCCATGCCATGTGTAATCCCAAGAGCTGTTACAAACTCAAGAAAGCAATGT	2858
Qy	781	AGATAATC--ATGCGTAAATTTGACGATGGGATCTGTAAATACGACTCACTATAGGGCA	838
Db	2857	GGTCTCTCTTTTGGTGGGATCTTTGAAAGGCGAGATTTGTGGACAGGTAATGGTGT	2798
Qy	839	ATTGGGTACCGGGCCCCCTCGAGGTCCAGCGTATCGATAAGCTTGATATCGAAATTCCT	898
Db	2797	CTGGTAAAGGACAGGGCCATCGCAATTTGGAGTATTTTGTGATAATGGTCTGTAGTT	2738
Qy	899	GCAGCCCGGGGATCCACTAGTTCTTAGAGCGCGCCACCGCGGTGGAGCTCCAGCTTTT	958
Db	2737	GAAAGCTTCCATCTTCAATGTGTGCTAAATTTTGAAGTTAACTTTGATTCATTTCTTTT	2678
Qy	959	GTTCCCTTTAGTGGGTTAATTAGATCCATGCGTCAATTTTACGACACTATCTTTCT	1018
Db	2677	GTTTGTGTCGCAAGATGATACATTTGTGTGA--GTTATAGTTGATTTCCAAATTTGTGTCA	2619
Qy	1019	AGGGTTAATCTAGCTGCATCAGGATCATATCGCGGTCTTTTTCGCGCTCAGTCATCG	1078
Db	2618	AGAAATGTTTCCATCTTCTTAAATCAATACCTTTTAATCGATTC----TATTAAACA	2564
Qy	1079	CCCAAGCTGGCGCTATCTGGGCAATCGGGAGGAAGAGCCCGTGTCTTTCCCGCGAGGT	1138
Db	2563	GGGTATCACTTCAAACTTGACTTC-----AGCACGTGTCTGTAGTTCCCGTC	2515
Qy	1139	TGAAGCGGATGGAAGAGTTTGGCGAGGATGACTGCTGCTCAATTTAGCTTGGCGAAA	1198
Db	2514	ATCCTTTGAAAATATAGTTCTTTCTGTATACATAACCTTCGGGCATGGCACTTTGAAANA	2455
Qy	1199	ACGCACGTTTACCATGATGATTTGGGAAAGGTGTGCCATGCACGCGCTTTTAAACGTTGA	1258
Db	2454	GTCAATGCCGTTTCAATATGATCTGGGTATCTTGAAAGCAATGAACACCATAGAGAAAT	2395
Qy	1259	GTTGTTACGGCCACTGGGATACAGTTCTGTGGCGCTTTTCCGGACACAGTTCCGGAT	1318
Db	2394	AGTGACAAAGTTGGCCCATGGAAACAGGTAGTTTCCAGTAGTGCATAAATAAATTTAAGGT	2335
Qy	1319	GGTCAGCCGAGCGATCAGCAACCCGNAACAATACCGCGGACAGCGGAACCTGCCGTGC	1378
Db	2334	AAGTTTTC-----CGTATGTTGCATCACCTTCACTCCCTCTCCACTGACAGAAATTTGTGC	2280
Qy	1379	CGGTGTGCAGATTAATGACAGCGGTGGCGCTGGGATATTACGTACAGCGAGGACGGGTA	1438
Db	2279	CCATTAAATCATCACCATCTAATTTCAACAGAAATGGGACAACTCCAGTGAAGTTCTTCT	2220
Qy	1439	TCCTGGCTGGATGCGCGAGAAATGGACATGGATACCCCGTAGATTACCCCGCGGGCGGC	1498
Db	2219	CTTTTACTCATCGTACCCAGCTTTTGTTCCTTTTGTAGTGAGGGTTA-----ATTGCGGC	2165
Qy	1499	TTGGCGTAAATCATGTTCATAGCTGTTTCTCGTGTGAATTTGTTATCCGCTCAACAATTTCA	1558
Db	2164	TTGGCGTAAATCATGTTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCAACAATTTCA	2105
Qy	1559	CACAACATACGAGCGGAGCATAAAGTGAAGCTTGGGTCGCTTAATGAGTGAGCTAA	1618
Db	2104	CACAACATACGAGCGGAGCATAAAGTGAAGCTTGGGTCGCTTAATGAGTGAGCTAA	2045
Qy	1619	CTCATTAAATTTGGTGGCTCTACCTGCCGCTTTCCAGTCGGGAAAACCTGTGTCGCGAG	1678
Db	2044	CTCATTAAATTTGGTGGCTCTACCTGCCGCTTTCCAGTCGGGAAAACCTGTGTCGTCGAG	1985
Qy	1679	CTGCATTAATTAATCGGCCAACCGCGGAGAGAGCGGTTTTCGCTATTGGGCGCTCTTCC	1738
Db	1984	CTGCATTAATTAATCGGCCAACCGCGGAGAGAGCGGTTTTCGCTATTGGGCGCTCTTCC	1925
Qy	1739	GCTTCTCGCTCACTGACTCGCTCGCTCGGTCGCTCGGTCGCGGAGCGGTATCAGCT	1798
Db	1924	GCTTCTCGCTCACTGACTCGCTCGCTCGGTCGCTCGGTCGCGGAGCGGTATCAGCT	1865
Qy	1799	CACTCAAAAGGCGGTAATACGGTTATCCACAGNATCAGGGGATAACGCGAGGAAGACATG	1858
Db	1864	CACTCAAAAGGCGGTAATACGGTTATCCACAGNATCAGGGGATAACGCGAGGAAGACATG	1805
Qy	1859	TCAGCAAAAGGCGGTAATACGGTTATCCACAGNATCAGGGGATAACGCGAGGAAGACATG	1918
Db	1804	TCAGCAAAAGGCGGTAATACGGTTATCCACAGNATCAGGGGATAACGCGAGGAAGACATG	1745
Qy	1919	CATAGGCTCCGCCCCCTTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCA	1978
Db	1744	CATAGGCTCCGCCCCCTTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCA	1685
Qy	1979	RACCCGACAGGACTATAAGATACCAAGGCTTCCCGCTGGAGCTCCCTCGTCGCTCT	2038
Db	1684	AACCCGACAGGACTATAAGATACCAAGGCTTCCCGCTGGAGCTCCCTCGTCGCTCT	1625
Qy	2039	CCTGTTCCGACCCCTTACCGGATACTGTCCGCTTCTCTCCCTTCGGGAAAGCGTG	2098
Db	1624	CCTGTTCCGACCCCTTACCGGATACTGTCCGCTTCTCTCCCTTCGGGAAAGCGTG	1565
Qy	2099	GGGCTTTCTATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTCGTCCCAAG	2158
Db	1564	GGGCTTTCTATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTCGTCCCAAG	1505
Qy	2159	CTGGGCTGTGCAGAACCCCGTTTCCAGCCGACCGCTGCTGCTTATCCCGTAACTAT	2218
Db	1504	CTGGGCTGTGCAGAACCCCGTTTCCAGCCGACCGCTGCTGCTTATCCCGTAACTAT	1445
Qy	2219	CGTCTTGAGTCCAAACCCGCTTAAGACACGACTTATCGCCACTGGCAGACCACTGGTAAC	2278
Db	1444	CGTCTTGAGTCCAAACCCGCTTAAGACACGACTTATCGCCACTGGCAGACCACTGGTAAC	1385
Qy	2279	AGGATTAGACAGCAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAAC	2338
Db	1384	AGGATTAGACAGCAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAAC	1325
Qy	2339	TACGGCTACCTAGAGGACAGTATTTGCTATCTCGCTCTGCTGACGACCACTTACCTTC	2398
Db	1324	TACGGCTACCTAGAGGACAGTATTTGCTATCTCGCTCTGCTGACGACCACTTACCTTC	1265
Qy	2399	GGAAAAAGATTGGTAGTCTTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGTTTT	2458
Db	1264	GGAAAAAGATTGGTAGTCTTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGTTTT	1205
Qy	2459	TTTGTTCGAGCAGCAGATTAACGCGCAAAAAGGATCTCAAGAGATCTTTTGTATC	2518
Db	1204	TTTGTTCGAGCAGCAGATTAACGCGCAAAAAGGATCTCAAGAGATCTTTTGTATC	1145
Qy	2519	TTTTCTACGCGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGGAATTTTGTCTAG	2578
Db	1144	TTTTCTACGCGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGGAATTTTGTCTAG	1085
Qy	2579	AGATTATCAAAAAGGATCTTCACTTAGATCTTTTAAATTTAAATTTAAATTTAAATCA	2638
Db	1084	AGATTATCAAAAAGGATCTTCACTTAGATCTTTTAAATTTAAATTTAAATTTAAATCA	1025
Qy	2639	ATCTAAAGTATATAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTAGGCA	2698
Db	1024	ATCTAAAGTATATAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTAGGCA	965
Qy	2699	CCTATCTCAGCGATCTGTCTATTTTCTGTTATCCATAGTTGCTGACTCCCGCTCGTGTAG	2758
Db	964	CCTATCTCAGCGATCTGTCTATTTCTGTTATCCATAGTTGCTGACTCCCGCTCGTGTAG	905

```
QY 2759 ATAACTACGATACGAGGCTTACCATCTGGCCCGAGTCTGCTCAATGATACCGCGAGAC 2818
Db 904 ATAACTACGATACGAGGCTTACCATCTGGCCCGAGTCTGCTCAATGATACCGCGAGAC 845
QY 2819 CCACGCTCACCAGCTCCAGATTATCAGCAATAAACCAGCAGCCGGAAGGGCCGAGCGC 2878
Db 844 CCACGCTCACCAGCTCCAGATTATCAGCAATAAACCAGCAGCCGGAAGGGCCGAGCGC 785
QY 2879 AGAAGTGGTCTGCTGCACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCCGGGAAGCT 2938
Db 784 AGAAGTGGTCTGCTGCACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCCGGGAAGCT 725
QY 2939 AGAGTAAGTAGTTGCCAGTAAATAGTTTGGCGCAAGTTGTGCCATTTGCTACAGGCATC 2998
Db 724 AGAGTAAGTAGTTGCCAGTAAATAGTTTGGCGCAAGTTGTGCCATTTGCTACAGGCATC 665
QY 2999 GTGGTGTACGCTCCGCTGTTGTTGATGCTTCAATTCAGCTCCGTTTCCCAACGATCAAG 3058
Db 664 GTGGTGTACGCTCCGCTGTTGTTGATGCTTCAATTCAGCTCCGTTTCCCAACGATCAAG 605
QY 3059 CGAGTTTACATGATCCCGCATGTTGTGCAAAAAAGCGTTAGCTCTCTCGTCTCTCGATC 3118
Db 604 CGAGTTTACATGATCCCGCATGTTGTGCAAAAAAGCGTTAGCTCTCTCGTCTCTCGATC 545
QY 3119 GTTGTGAGTAAGTTGGCCGAGTGTATCACTCATGTTTATGCGCAGCAGTCAATAT 3178
Db 544 GTTGTGAGTAAGTTGGCCGAGTGTATCACTCATGTTTATGCGCAGCAGTCAATAT 485
QY 3179 TCTCTTACTGTCAATGCTCCGTAAGATGTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3238
Db 484 TCTCTTACTGTCAATGCTCCGTAAGATGTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 425
QY 3239 TCATTTCTGAGATAGTGTATGCGCGACCGAGTTGCTCTTTCGCGCGCTGCTCAATACCGGAT 3298
Db 424 TCATTTCTGAGATAGTGTATGCGCGACCGAGTTGCTCTTTCGCGCGCTGCTCAATACCGGAT 365
QY 3299 AATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAAAAGTTCTTCGGGG 3358
Db 364 AATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAAAAGTTCTTCGGGG 305
QY 3359 CGAAACTCTCAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCCTCTGTCA 3418
Db 304 CGAAACTCTCAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCCTCTGTCA 245
QY 3419 CCCAACTGATCTTCAGCATCTTTTACTTTCCACGAGGTTTCTGGGTGAGCAAAAAAGGA 3478
Db 244 CCCAACTGATCTTCAGCATCTTTTACTTTCCACGAGGTTTCTGGGTGAGCAAAAAAGGA 185
QY 3479 AGGCAAAATGCCGCAAAAAAGGAATAAGGGCGACACGGAATGTTGTAATCTCATCTC 3538
Db 184 AGGCAAAATGCCGCAAAAAAGGAATAAGGGCGACACGGAATGTTGTAATCTCATCTC 125
QY 3539 TTCCCTTTTCAATATTTAAGCATTTATCAGGTTATGTTCTCATGAGCGGATACATA 3598
Db 124 TTCCCTTTTCAATATTTAAGCATTTATCAGGTTATGTTCTCATGAGCGGATACATA 65
QY 3599 TTTGAATGTATTTAGAAAAATAAACAATAAGGGGTTCCGCGCACATTTCCCGAAAAAGTG 3658
Db 64 TTTGAATGTATTTAGAAAAATAAACAATAAGGGGTTCCGCGCACATTTCCCGAAAAAGTG 5
QY 3659 CCAC 3662
Db 4 CCAC 1
```

RESULT 14

ADQ88437/C

ID ADQ88437 standard; DNA; 3637 BP.

XX AC

AC ADQ88437;

XX AC

DT 18-NOV-2004 (first entry)

XX XX

```
DE PBSWTGFP DNA encoding Jellyfish wild-type GFP protein.
XX Molecular biology; jellyfish; green fluorescent protein; GFP; ds.
KW Aequorea victoria.
OS US2004142433-A1.
PN 22-JUL-2004.
XX 10-OCT-2003; 2003US-00684134.
XX 02-FEB-2001; 2001US-0266386P.
PR 14-FEB-2001; 2001US-0268785P.
PR 01-FEB-2002; 2002US-00066390.
PR 08-AUG-2002; 2002US-0402342P.
PR 21-AUG-2002; 2002US-00226372.
PR 25-OCT-2002; 2002US-00280913.
PR 08-AUG-2003; 2003US-00637758.
XX (PADG/) PADGETT H S.
PA (FITZ/) FITZMAURICE W P.
PA (LIND/) LINDBO J A.
PA (VAEW/) VAETHONGS A A.
PA (VOJD/) VOJDANI F S.
PA (SMIT/) SMITH M L.
XX Padgett HS, Fitzmaurice WP, Lindbo JA, Vaethongs AA, Vojdani FS;
PI Smith ML;
XX WPI; 2004-552565/53.
XX Preparing variant polynucleotides having different nucleotide sequences
PT from at least two parent polynucleotides, useful in molecular biology, in
PT particular for generating populations of related nucleic acid molecules.
XX Example 5; SEQ ID NO 31; 81pp; English.
XX The invention relates to a method of preparing a variant polynucleotide
CC having a different nucleotide sequence from at least two parent
CC polynucleotides. The method involves preparing at least one heteroduplex
CC between the two parent polynucleotides, cleaving at least one
CC polynucleotide strand in the heteroduplex at a mismatch site to form a
CC cleavage site, replacing at least one nucleotide on at least one strand
CC at or near the cleavage site where at least one of the at least two parent
CC different nucleotide sequence from either of the at least two parent
CC polynucleotides. Methods and compositions of the invention are useful in
CC the field of molecular biology in particular for generating populations
CC of related nucleic acid molecules. The present sequence is pBSWTGFP.
CC plasmid DNA encoding Jellyfish wild-type GFP (green fluorescent protein).
CC This sequence is used in the exemplification of the invention
XX
SQ Sequence 3637 BP; 959 A; 879 C; 868 G; 931 T; 0 U; 0 Other;
Query Match 72.5%; Score 2656.4; DB 13; Length 3637;
Best Local Similarity 84.0%; Pred. NO. 0;
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;
QY 1 CTAAATTGTAAGCGTTAATATTTTGTAAAATTTCGCGTTAAATTTTGTAAATCAGCTC 60
Db 3637 CTAAATTGTAAGCGTTAATATTTTGTAAAATTTCGCGTTAAATTTTGTAAATCAGCTC 3578
QY 61 ATTTTAAACCAATAGCGCAAAATCGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 120
Db 3577 ATTTTAAACCAATAGCGCAAAATCGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 3518
QY 121 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTGACTC 180
Db 3517 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTGACTC 3458
QY 181 CAACGTCAAGGGCGGAAACCCGCTTATCAGGGCGATGCCCCACTACGTGAACCATCACC 240
Db 3457 CAACGTCAAGGGCGGAAACCCGCTTATCAGGGCGATGCCCCACTACGTGAACCATCACC 3398
```

QY 241 CTAATCAAGTTTTTTGGGCTCGAGGTGCGGTAAAGCACTAAATCGGAACCCCTAAAGGGAG 300
Db |||||
QY 3397 CTAATCAAGTTTTTTGGGCTCGAGGTGCGGTAAAGCACTAAATCGGAACCCCTAAAGGGAG 3338
Db |||||
QY 301 CCCCCGATTTAGAGCTTGAACGGGGAAGCCGCGGAACGTTGGCGAGAAAGGAAGGAAGAA 360
Db |||||
QY 3337 CCCCCGATTTAGAGCTTGAACGGGGAAGCCGCGGAACGTTGGCGAGAAAGGAAGGAAGAA 3278
Db |||||
QY 361 AGCGAAGCAGCGGGCGCTAGGGCGCTGCAAGTGTAGCGGTACGCTCGCGGTAAACCA 420
Db |||||
QY 3277 AGCGAAGCAGCGGGCGCTAGGGCGCTGCAAGTGTAGCGGTACGCTCGCGGTAAACCA 3218
Db |||||
QY 421 CACACCCGCGCGCTTAATGCGCCGCTACAGCGCGCGTCCCATTTGCCATTTCAGGCTGCG 480
Db |||||
QY 3217 CACACCCGCGCGCTTAATGCGCCGCTACAGCGCGCGTCCCATTTGCCATTTCAGGCTGCG 3158
Db |||||
QY 481 CAACTGTTGGGAAGGGCGATTCGGTGGGCGCTTTCGCTATTACGCGAGCTGGCGAAAGG 540
Db |||||
QY 3157 CAACTGTTGGGAAGGGCGATTCGGTGGGCGCTTTCGCTATTACGCGAGCTGGCGAAAGG 3098
Db |||||
QY 3097 GGGATGTGCTCAAGCGCAATTAAGTTGGGTAAAGTGGGTAAAGTGGGTAAAGTGGGTAAAG 3038
Db |||||
QY 601 TAAAAACGCGCAGTACGCGCGCTCGTTTCATTTCACGTTTTCGAAACCCGTCGGAGACGG 660
Db |||||
QY 3037 TAAAAACGCGCAGTACGCGCGCTCGTTTCATTTCACGTTTTCGAAACCCGTCGGAGTCCA 2978
Db |||||
QY 661 GCAGACTCGCGTGCAGAAATGTTTTTACAGCGTGTAGGAGCAGATGAAGATGCTCGACAC 720
Db |||||
QY 2977 CCGCGTGGCGCGCGCTCTAGAACTAGTGTGATCCCGCGGCTGCAGGAATCTTATTGTT 2918
Db |||||
QY 721 GCTGCAGACACGCGAGCTAGATTAAACCTAGAAAGATATATATGTCAGCTACGTTAA 780
Db |||||
QY 2917 ATAGTTTCATCCATGCCATGTGTAATCCCGAGCGTGTTCACAACTCAAGAAGGACCATGT 2858
Db |||||
QY 781 AGATAATC--ATGCGGTAAATTAAGCGCATGCTGTAATACGACTCACTATAGGGCGCA 838
Db |||||
QY 2857 GGTCTCTCTTTTCGTTGGGATCTTCGAAAGGCGAGATGTTGTGACAGGTAATGTTGT 2798
Db |||||
QY 839 ATTGGGTACCGGGCCCCCTCGAGGTGCAACGGTATCGATTAAGCTTGATATCGAATTCCT 898
Db |||||
QY 2797 CTGGTAAAGGACAGGGCCATCGCCAAATGGAGTATTTTGTGATAATGGTCTGTAGTT 2738
Db |||||
QY 899 GCAGCCCGGGGATCCACTAGTTCTAGAGCGCCGCCACCGGCTGGAGCTCCAGCTTTT 958
Db |||||
QY 2737 GAACGCTTCCATCTTCAATGTTGTGTTAAATTTTGAAGTTAACTTTGATTCCATTCTTTT 2678
Db |||||
QY 959 GTTCCCTTTTAGTGAGGGTTAAATTAGATCCATCGCGTCAATTTTACGCACTATCTTTCT 1018
Db |||||
QY 2677 GTTTGCTGCCATGATATACATTTGTGTGA--GTTATAGTTGATTTCCAAATTTGTGTCCA 2619
Db |||||
QY 1019 AGGGTTAATCTAGCTGCAATCAGGATCATATCTCGGGGTCTTTTTCGGGCTCAGTCAATCG 1078
Db |||||
QY 2618 AGAATGTTTCCATCTCTTTTAAATCAATACCTTTTAACTCGATTC--TATTAAACAA 2564
Db |||||
QY 1079 CCCAAGCTGGGCTATCTGGGCTACGGGAGGAAGACCGGTGCTTTTCCCGGAGGT 1138
Db |||||
QY 2563 GGGTATCACCTTCAACTTGACTTC-----AGCACGTGCTTGTAGTCCCGTC 2515
Db |||||
QY 1139 TGAAGCGGCATGGAAGAGTTTTCGCGAGGATGACTGCTGCTGCAATGAGTTGAGCGGAAA 1198
Db |||||
QY 2514 ATCTTTGAAAATATAGTTCTTTCTGCTGATACAACTTCGGGCATGGCACTCTTGAANA 2455
Db |||||
QY 1199 AGCAGGTTTACATGATGATTCGGGAAGGTTGGGCATGCAACGCTTTTAAACGTTGAAT 1258
Db |||||
QY 2454 GTCATGCGGTTTTCATATGATCTGGGTATCTTGAAGAAGCATTTGAACCATGAAGAAAGT 2395
Db |||||
QY 1259 GTTTCGTTCAGGCCACTCGGATACCAATTCGTCGGGCTTTTTCGGACACAGTTCGGAT 1318
Db |||||
QY 2394 AGTGACAAAGTTTGCCATGGAAACAGGTAGTTTTCAGTAGTGTGCAATAAATTTAAGGTT 2335
Db |||||

QY 1319 GGTACAGCCGAAGCGCATCAGCAACCCGAAACAATATCCGCGCAGCAGCCGGAACCTCGCGTGC 1378
Db |||||
QY 2334 AAGTTTTTC-----CGTATGTTGCATCACCTTCACTTCACTGACAGAAAAATTTGTGC 2280
Db |||||
QY 1379 CGGTGTGCAGATTAAATGACAGCGGTGCGGCTGGGATATTAAGTTCAGTTCAGCGAGACGGGTA 1438
Db |||||
QY 2279 CCAATTAACATCACCATCTAAATTCACAAGAAATGGGACAACTCCAGTGAAGAGTTCTTCT 2220
Db |||||
QY 1439 TCTCGCTGGATGCGCGAGAAATGACATGATATCCCGTGAAGTACCCGCGGGCGCGC 1498
Db |||||
QY 2219 CTTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAAGGTTA-----ATTGGCGC 2165
Db |||||
QY 1499 TTGGCGTAATCATGTCTAGCTGTTCTCTGTGTGAAATTTGTTATCCGCTCAATTCOA 1558
Db |||||
QY 2164 TTGGCGTAATCATGTCTAGCTGTTCTCTGTGTGAAATTTGTTATCCGCTCAATTCOA 2105
Db |||||
QY 1559 CACAACATACGAGCGCGAAGCATAAAGTGTAAAGCTGGGTGCCTTAATGAGTGAAGTAA 1618
Db |||||
QY 2104 CACAACATACGAGCGCGAAGCATAAAGTGTAAAGCTGGGTGCCTTAATGAGTGAAGTAA 2045
Db |||||
QY 1619 CTCACATTAATTTGCGTTGCGCTCACTGCGCGCTTTCAGTTCGGGAAACCTGTGTCGCCAG 1678
Db |||||
QY 2044 CTCACATTAATTTGCGTTGCGCTCACTGCGCGCTTTCAGTTCGGGAAACCTGTGTCGCCAG 1985
Db |||||
QY 1679 CTGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTCGCTATTGGGCGCTCTTCC 1738
Db |||||
QY 1984 CTGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTCGCTATTGGGCGCTCTTCC 1925
Db |||||
QY 1739 GCTTCTCGTCACTGACTCGCTGCGCTCGTGTGTTTCGCTTCGGGAGCGGTATTCAGT 1798
Db |||||
QY 1924 GCTTCTCGTCACTGACTCGCTGCGCTCGTGTGTTTCGCTTCGGGAGCGGTATTCAGT 1865
Db |||||
QY 1799 CACTCAAAAGGGCGTAATACGCTTATCCAGAAATCAGGGGATAACGAGAAAGAACATG 1858
Db |||||
QY 1864 CACTCAAAAGGGCGTAATACGCTTATCCAGAAATCAGGGGATAACGAGAAAGAACATG 1805
Db |||||
QY 1859 TGAGCAAAAGGGCGCAAAAGGCCAGGAAACCGTAAAAAGGCGCGTTGCTGCGCTTTTC 1918
Db |||||
QY 1804 TGAGCAAAAGGGCGCAAAAGGCCAGGAAACCGTAAAAAGGCGCGTTGCTGCGCTTTTC 1745
Db |||||
QY 1919 CATAGGCTTCGCGCCCTTGAACGAGCATCACAAAATTCAGCGCTCAAGTCAGAGGTGGGA 1978
Db |||||
QY 1744 CATAGGCTTCGCGCCCTTGAACGAGCATCACAAAATTCAGCGCTCAAGTCAGAGGTGGGA 1685
Db |||||
QY 1979 AACCCGACAGGACTATAAGATACCAAGCGTTTCCTCCCTGGAAGCTCCCTCGTGGCTCT 2038
Db |||||
QY 1684 AACCCGACAGGACTATAAGATACCAAGCGTTTCCTCCCTGGAAGCTCCCTCGTGGCTCT 1625
Db |||||
QY 2039 CTTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTG 2098
Db |||||
QY 1624 CTTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTG 1565
Db |||||
QY 2099 GGGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGCGGTAGGTGCTGCTGCCAG 2158
Db |||||
QY 1564 GGGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGCGGTAGGTGCTGCTGCCAG 1505
Db |||||
QY 2159 CTGGGCTGTGTGACGAACCCCGCTTACGCGCGACCGCTGCGCTTATTCGGTAACTAT 2218
Db |||||
QY 1504 CTGGGCTGTGTGACGAACCCCGCTTACGCGCGACCGCTGCGCTTATTCGGTAACTAT 1445
Db |||||
QY 2219 GGTCTTGAAGTCCAAACCCCGTAAAGACACGACTTATCGCACTGGCAGACCACTGTAAC 2278
Db |||||
QY 2279 AGGATTACGAGCGAGGTATGTAGGCGGTCTACAGAGTTCTTCAAGTGTGGGCTAAC 2338
Db |||||
QY 1384 AGGATTACGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGGCTAAC 1325
Db |||||
QY 2339 TACGGCTACCTAGAAAGACAGTATTTTGGTATCTCGCTCTGCTGAAGCCAGTTTACCTTC 2398
Db |||||
QY 1324 TACGGCTACCTAGAAAGACAGTATTTTGGTATCTCGCTCTGCTGAAGCCAGTTTACCTTC 1265
Db |||||
QY 2399 GGAAGAAAGAGTTGGTGTAGCTCTTGATTCGGGCAAAACAAACCGCTGGTAGCGGTGTTTT 2458
Db |||||

1264	Db		GGAAAAAGAGTTCGGTAGCTCTTGATCCGGCAAA	CAAA	CCACCGCTGGTAGCGGTGGTTTT	1205
2459	Qy	TTTTTTT	TGCAAGCAGCAGATTTACGCGCAGAAAAA	AGGATCTCAAGAAGATCTCTTTGATC	2518	
1204	Db	TTTTTTT	TGCAAGCAGCAGATTTACGCGCAGAAAAA	AGGATCTCAAGAGATCTCTTTGATC	1145	
2519	Qy	TTTTTCTA	CGGGGTCTGACGCTCAGTGGAA	CGAAAACTCAGTTTAAGGATTTTTGGTCATG	2578	
1144	Db	TTTTTCTAC	GGGTCTGACGCTCAGTGGAA	CGAAAACTCAGTTTAAGGATTTTTGGTCATG	1085	
2579	Qy	AGATTATCA	AAAAAGGATCTTCACTAGATCTTTTAAAT	TTAAAAATGAAGTTTAAATCA	2638	
1084	Db	AGATTATCA	AAAAAGGATCTTCACTAGATCTTTTAAAT	TTAAAAATGAAGTTTAAATCA	1025	
2639	Qy	ATCTAAAGTATAT	TAGTAAATCTTGCTCTGACAGTTTACCAAT	TGCTTAATCAGTAGGCA	2698	
1024	Db	ATCTAAAGTATAT	TAGTAAATCTTGCTCTGACAGTTTACCAAT	TGCTTAATCAGTAGGCA	965	
2699	Qy	CCATATCTCA	CGCATCTGTCTATTTTCGTTTCATC	ATAGTTGCCTGACTCCCGCTCGTGTAG	2758	
964	Db	CCATATCTCA	CGCATCTGTCTATTTTCGTTTCATC	ATAGTTGCCTGACTCCCGCTCGTGTAG	905	
2759	Qy	ATAACTACGATAC	GGAGGGCTTACCATCTGGGCCAGTCTGCAAT	TGATACCGCAGAC	2818	
904	Db	ATAACTACGATAC	GGAGGGCTTACCATCTGGGCCAGTCTGCAAT	TGATACCGCAGAC	845	
2819	Qy	CCACGCTCA	CCGGCTCCAGATTTATCAGCAAT	TAAACACGACCGGAGGGCCGAGCGC	2878	
844	Db	CCACGCTCA	CCGGCTCCAGATTTATCAGCAAT	TAAACACGACCGGAGGGCCGAGCGC	785	
2879	Qy	AGAAAGTGGTCT	GCACCTTTATCCGCTCCATCCAGTCTATTAAT	TGTTGCCGGGAAGCT	2938	
784	Db	AGAAAGTGGTCT	GCACCTTTATCCGCTCCATCCAGTCTATTAAT	TGTTGCCGGGAAGCT	725	
2939	Qy	AGAGTAAGTATG	TGCGCAGTTAATAGTTTGC	GCAACGTTTGCCATCAGGCATC	2998	
724	Db	AGAGTAAGTATG	TGCGCAGTTAATAGTTTGC	GCAACGTTTGCCATCAGGCATC	665	
2999	Qy	GTGGTGTCA	CGCTCGTGTGTTGATGGCTTCAATCAGCT	CCGGTCCCAACCATCAAGG	3058	
664	Db	GTGGTGTCA	CGCTCGTGTGTTGATGGCTTCAATCAGCT	CCGGTCCCAACCATCAAGG	605	
3059	Qy	CGAGTTACATGAT	CCCCCATCTTGTGCAAAAAAGCGTTAGCT	CTCTCGGTCTCCGATC	3118	
604	Db	CGAGTTACATGAT	CCCCCATCTTGTGCAAAAAAGCGTTAGCT	CTCTCGGTCTCCGATC	545	
3119	Qy	GTGTGAGAA	GTGTTGGCCGAGTGTTATCACTCATG	TTATGGCAGCACTGCATAT	3178	
544	Db	GTGTGAGAA	GTGTTGGCCGAGTGTTATCACTCATG	TTATGGCAGCACTGCATAT	485	
3179	Qy	TCTCTTACTGT	CATGCCATCCGTAAGTGTCTTCTG	TGACTGGTAGTACTCAACGAG	3238	
484	Db	TCTCTTACTGT	CATGCCATCCGTAAGTGTCTTCTG	TGACTGGTAGTACTCAACGAG	425	
3239	Qy	TCATTTGAGAA	TAGTGTATGCGCGACCGAGTTGCTCTT	GCCCGCGTCAATACGGGAT	3298	
424	Db	TCATTTGAGAA	TAGTGTATGCGCGACCGAGTTGCTCTT	GCCCGCGTCAATACGGGAT	365	
3299	Qy	AATACCGCG	CCACATAGCAGAACTTTTAAAGTGCTCAT	CAATTCGAAAAACGTTCTTCGGGG	3358	
364	Db	AATACCGCG	CCACATAGCAGAACTTTTAAAGTGCTCAT	CAATTCGAAAAACGTTCTTCGGGG	305	
3359	Qy	CGAAAACTC	TAAGGATCTTACCGCTGTGTAGATCCAGT	TGTAACCCACTCTGTGCA	3418	
304	Db	CGAAAACTC	TAAGGATCTTACCGCTGTGTAGATCCAGT	TGTAACCCACTCTGTGCA	245	
3419	Qy	CCCAACTGAT	CTTTCAGCATCTTTTACTTTTCCACAG	GGTTTTCTGGGTAGCAAAAAACAGGA	3478	
244	Db	CCCAACTGAT	CTTTCAGCATCTTTTACTTTTCCACAG	GGTTTTCTGGGTAGCAAAAAACAGGA	185	
3479	Qy	AGGCAAAAT	CGCCAAAAAAGGAAATAGGGCGCA	CACGAAATGTGTAATCTCATCTC	3538	

Db	184	AGGCAAAATCCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATCTACTC	125
Qy	3539	TTCTCTTTTCAATATATTATGAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATA	3598
Db	124	TTCTCTTTTCAATATATTATGAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATA	65
Qy	3599	TTTGAATGATTTTAGAAAAATAAACAATAAGGGTTCCGGCGACATTTCCCCGAAAAGTG	3658
Db	64	TTTGAATGATTTTAGAAAAATAAACAATAAGGGTTCCGGCGACATTTCCCCGAAAAGTG	5
Qy	3659	CCAC 3662	
Db	4	CCAC 1	
RESULT 15			
ADSL17191/c			
ID	ADSL17191	standard; DNA; 3637 BP.	
XX	AC		
XX	ADSL17191;		
DT	02-DEC-2004	(first entry)	
XX	pSWTGFPP	plasmid DNA encoding A. victoria GFP Cycle 3 ORF.	
DE	Polymerase; GFP;	jellyfish; green fluorescent protein; chimeric; gene;	
XX	ds.		
KW	Aequorea victoria.		
OS	Chimeric.		
OS	Unidentified.		
XX			
PN	US2004180352-A1.		
XX			
PD	16-SEP-2004.		
XX			
PF	08-AUG-2003; 2003US-00637758.		
XX			
PR	08-AUG-2002; 2002US-0402342P.		
XX			
PA	(LARG-) LARGE SCALE BIOLOGY CORP.		
XX			
PI	Padgett HS, Lindbo JA, Fitzmaurice WP, Vaezhongs AA;		
XX	WPI; 2004-667658/65.		
DR			
XX			
PT	In vitro method of making sequence variants from heteroduplex		
PT	polynucleotide, involves combining polynucleotide with agent having		
PT	polymerase and strand cleavage activities for sufficient time for		
PT	percentage of complementarity to increase.		
XX			
PS	Example 18; SEQ ID NO 31; 81pp; English.		
XX			
CC	The invention relates to an in vitro method of making sequence variants		
CC	from heteroduplex polynucleotide which involves combining polynucleotide		
CC	with an agent having polymerase and strand cleavage activities for		
CC	sufficient time for the percentage of complementarity to be increased		
CC	within the heteroduplex. The method is useful for performing in vitro		
CC	method of making sequence variants from one or more heteroduplex		
CC	polynucleotide. It is useful in increasing diversity in a population of		
CC	sequences and obtaining a polynucleotide encoding a desired functional		
CC	property. The present sequence is a plasmid DNA encoding Aequorea		
CC	victoria green fluorescent protein (GFP) Cycle' 3 ORF DNA. This sequence		
CC	is used to illustrate the method of the invention.		
XX			
SQ	Sequence 3637 BP; 959 A; 879 C; 868 G; 931 T; 0 U; 0 Other;		
Query Match 72.5%; Score 2656.4; DB 13; Length 3637;			
Best Local Similarity 84.0%; Pred. No. 0;			
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;			
Qy	1	CTAAATGTTAAGCGTTAATAATTTTGTAAATTCGGGTAAATTTTGTAAATCAAGTTC	60

Dd 3637 CTAATTTGTAAGCGTTAAATATATTTTGTAAAAATTCGCGTTAAATTTTGTAAAAATCAGCTC 3578
Qy 61 ATTTTAAACCAATAGGCGCAAAATCGCAAAATCCCTTATTAATCAAAAGAAATAGACGA 120
Dd |||||
Dd 3577 ATTTTAAACCAATAGGCGCAAAATCGCAAAATCCCTTATTAATCAAAAGAAATAGACGA 3518
Qy 121 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAAAGTGGACTC 180
Dd |||||
Dd 3517 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAAAGTGGACTC 3458
Qy 181 CAACTCAAGGCGGAAACCGTCTATCAGGGCGATGCCCCACTACGTGAAACCAATCAC 240
Dd |||||
Dd 3457 CAACTCAAGGCGGAAACCGTCTATCAGGGCGATGCCCCACTACGTGAAACCAATCAC 3398
Qy 241 CTAATCAAGTTTGGGTCGAGTCCGTAAGCACTAAATCGGAACCCCTAAAGGGAG 300
Dd |||||
Dd 3397 CTAATCAAGTTTGGGTCGAGTCCGTAAGCACTAAATCGGAACCCCTAAAGGGAG 3338
Qy 301 CCCCAGATTAGAGCTTGAACGGGGAAGCGCGCAACGTGGGAGAAAGGAGGAAAGAA 360
Dd |||||
Dd 3337 CCCCAGATTAGAGCTTGAACGGGGAAGCGCGCAACGTGGGAGAAAGGAGGAGAA 3278
Qy 361 AGCGAAAGAGCGGCGCTAGGCGGTGCAAGTGTAGCGGTCAAGTGTAGCGGTAAACAC 420
Dd |||||
Dd 3277 AGCGAAAGAGCGGCGCTAGGCGGTGCAAGTGTAGCGGTCAAGTGTAGCGGTAAACAC 3218
Qy 421 CACACCGCGCGCTTAATCGCCGCTACAGGCGCGTCCCAATCGCCATTCAGGCTGCG 480
Dd |||||
Dd 3217 CACACCGCGCGCTTAATCGCCGCTACAGGCGCGTCCCAATCGCCATTCAGGCTGCG 3158
Qy 481 CAACTGTTGGGAAGGCGCATCGGTGCGGCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
Dd |||||
Dd 3157 CAACTGTTGGGAAGGCGCATCGGTGCGGCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098
Qy 541 GGGATGTGCTCAAGGCGAATTAAGTTGGGTAAAGCGGAGGTTTCCAGTCAAGCGTTG 600
Dd |||||
Dd 3097 GGGATGTGCTCAAGGCGAATTAAGTTGGGTAAAGCGGAGGTTTCCAGTCAAGCGTTG 3038
Qy 601 TAAACGACGCGCAGTAGCGCGCTCTGTTCAATTCAGCTTTTGAACCGTGAGGACGG 660
Dd |||||
Dd 3037 TAAACGACGCGCAGTAGCGCGCTCTGTTCAATTCAGCTTTTGAACCGTGAGGACGG 2978
Qy 661 GCAGACTCGCGTCAAAATGTTTTACAGCGTGTATGGAGCAGATGAAGATGCTCGACAC 720
Dd |||||
Dd 2977 CCGCGTGGGCGCGCTCTAGAACTAGTGTGATCCCGCGCTGCAGGAATCTTATTGT 2918
Qy 721 GCTGCAGAACGCGAGCTAGATTAAACCTAGAAAGATTAATCATATTGTGACGTAGTTAA 780
Dd |||||
Dd 2917 ATAGTTTATCCATGCCATGTGTAATCCCGACAGCTGTTACAACTCAAGAGGACCATGT 2858
Qy 781 AGATAATC--ATGCGTAAATTTGACGATGGATCTGTAATACGACTCACTATAGGCGA 838
Dd |||||
Dd 2857 GGTCTCTCTTTTGTGGGATCTTTGAAAGGCGAGATTGTGTGACAGGTAATGGTGT 2798
Qy 839 ATTGGGTACCGGCGCCCGCTCGAGGTGCAGCGTATCGATAAGCTTGATATCGAATTCCT 898
Dd |||||
Dd 2797 CTGGTAAAGGACAGGCGCATCGCAATTGGAGTATTTGTGATAATGGTCTGCTAGTT 2738
Qy 899 GCAGCGGGGGATCCACTAGTTCTTAGAGCGCGCCCGCCCGCGGTGGAGCTCCAGCTTTT 958
Dd |||||
Dd 2737 GAAACGCTTCCATCTTCAATGTTGTGCTAAATTTTGAAGTTAACTTTGATTCATTTCTTTT 2678
Qy 959 GTTCCCTTTTAGTGGGTTAATTAGATCCATGCGTCAATTTTACGACAGCTATCTTTCT 1018
Dd |||||
Dd 2677 GTTGTCTGCCATGATATACATTGTGTGA--GTATAGTTGTATTTCCAAATTGTGTCA 2619
Qy 1019 AGGGTAAATCTAGCTGCATCAGGATCATATCGCGGTCTTTTTTCCGCTCAGTCATCG 1078
Dd |||||
Dd 2618 AGAATGTTTCCATCTTCTTAAATCAATACCTTTTAACTCGATTCTC---TATTAA 2564
Qy 1079 CCCAAGCTGGCGCTATCTGGGATCGGGAGGAAAGCGCGTGTCTTTTCCCGGAGGT 1138
Dd |||||
Dd 2563 GGGTATCACCTTCAAACTTGACTTC-----AGCACGTGTCTTTGTAGTTCCCGTC 2515

Qy 1139 TGAACGGCATGGAAGAGTTTTCGAGGATGACTGCTGCTGACTTGAAGTTGAGCGAAA 1198
Dd |||||
Dd 2514 ATCCTTGAATAATATAGTTCTTCTGTACATAACCTTCGGGCATGCGACTCTTGAANA 2455
Qy 1199 ACGCACGTTTACATGATGATTCGGAAAGGTGTGCCATGCAAGCTTTTAAACGGTGAAC 1258
Dd |||||
Dd 2454 GTCATGCCGTTTCAATATGATCTGGGTATCTTTGAAAGCATTCGAACACCATAGAGAAAG 2395
Qy 1259 GTTCTGTCAGGCCACTGGGATACAGTTCTGTCGGGCTTTTCCGGACACAGTTCCGGAT 1318
Dd |||||
Dd 2394 AGTGAACAAGTGTGGCATGGAACAGGTAGTTTCCAGTAGTGCAATAAATTTAAGG 2335
Qy 1319 GGTCAAGCCGAAGCGCATCAGCAACCGAAACAATACCGCGACACCGCGAACTCGCGTC 1378
Dd |||||
Dd 2334 AAGTTTTC-----CGTATGTTGCATCACCTTCCACTCAGACGAAATTTGTG 2280
Qy 1379 CGGTGTGAGATTAATGACAGCGGTGCGCGCTGGGATATTAATCGTACAGGAGCGGTA 1438
Dd |||||
Dd 2279 CCATTAACATCACCATCTAATTTCAACAAGATTGGGACAACTCCAGTGAAGATTTCTCT 2220
Qy 1439 TCCTGGCTGGATGCCGCGAGAAATGACACATGATACCCGCTGAGTTACC CGCGCGCGCG 1498
Dd |||||
Dd 2219 CCTTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTA-----ATTGCGCG 2165
Qy 1499 TTGGCGTAATCATGCTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTTCCA 1558
Dd |||||
Dd 2164 TTGGCGTAATCATGCTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTTCCA 2105
Qy 1559 CACAACTACGAGCGGAGCATAAAGTGTAAAGCTCGGGTGCCTAATAGTAGAGCTAA 1618
Dd |||||
Dd 2104 CACAACTACGAGCGGAGCATAAAGTGTAAAGCTCGGGTGCCTAATAGTAGAGCTAA 2045
Qy 1619 CTCACATTAATTGCGTTGCGCTCACTGCGCTTTTCCAGTCCGGAACCTGTGCTGCCAG 1678
Dd |||||
Dd 2044 CTCACATTAATTGCGTTGCGCTCACTGCGCTTTTCCAGTCCGGAACCTGTGCTGCCAG 1985
Qy 1679 CTGCATTAATGAATCGGCCAAACCGCGCGGAGAGCGGTTTTCGTTATTTGGGCGCTTCTCC 1738
Dd |||||
Dd 1984 CTGCATTAATGAATCGGCCAAACCGCGCGGAGAGCGGTTTTCGTTATTTGGGCGCTTCTCC 1925
Qy 1739 GCTTCTCGCTCACTGACTCGCTGCGCTGCTGCTGCTGCTGCGCGAGCGGTATCAGCT 1798
Dd |||||
Dd 1924 GCTTCTCGCTCACTGACTCGCTGCGCTGCTGCTGCTGCGCGAGCGGTATCAGCT 1865
Qy 1799 CACTCAAGGCGGTAAATACGTTATCCACAGATCAGGCGATTAACGAGGAAGAACATG 1858
Dd |||||
Dd 1864 CACTCAAGGCGGTAAATACGTTATCCACAGATCAGGCGATTAACGAGGAAGAACATG 1805
Qy 1859 TGAGCAAAAGGCGCAGCAAAAGCGCAGGAACCGTAAAGGCGCGTGTGCGCGCTTTTTC 1918
Dd |||||
Dd 1804 TGAGCAAAAGGCGCAGCAAAAGCGCAGGAACCGTAAAGGCGCGTGTGCTGCGCTTTTC 1745
Qy 1919 CATAGGCTCCGCGCGCTGACGAGCATCAAAAAATCAACGCTCAAGTCAGAGGTGGCA 1978
Dd |||||
Dd 1744 CATAGGCTCCGCGCGCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCA 1685
Qy 1979 AACCCGACGAGCTATAAGATACAGGCTTTCCCTCGGAAGCTCCCTCGTGGCTCT 2038
Dd |||||
Dd 1684 AACCCGACGAGCTATAAGATACAGGCTTTCCCTCGGAAGCTCCCTCGTGGCTCT 1625
Qy 2039 CTTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAAGCGTG 2098
Dd |||||
Dd 1624 CTTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTG 1565
Qy 2099 GCGCTTTCTATAGCTACGCTCAGGTAGGTATCTCAGTTGCTGTTAGGTGTTGCTCCCAAG 2158
Dd |||||
Dd 1564 GCGCTTTCTATAGCTACGCTCAGGTAGGTATCTCAGTTGCTGTTAGGTGCTTCCGCTCCAA 1505
Qy 2159 CTGGCTGTGTGACGAACCCCGTTCCAGCCGACCGCTGCGCTTATCCGTTAACTAT 2218
Dd |||||
Dd 1504 CTGGCTGTGTGACGAACCCCGCTTCCAGCCGACCGCTGCGCTTATCCGTTAACTAT 1445

[illegible]

Search completed: June 5, 2005, 14:16:59
Job time : 1798.75 secs

364	Db	AATACCGCGCCATAGCAGAACTTTAAAAAGTGTCTATCATGTGAAACGTTCTTCTCGGGG	305
3359	Qy	CGAAAACTCTCAAGGATCTTTACCGCTGTTCAGATCCAGTTTCGATGTAAACCACTCTGNGCA	3418
304	Db	CGAAAACTCTCAAGGATCTTTACCGCTGTTCAGATCCAGTTTCGATGTAAACCACTCTGNGCA	245
3419	Qy	CCAACTGTATCTTCAGCATCTTTTACTTTTCACACAGCGTTTCTGGGTGAGCAAAAACAGGA	3478
244	Db	CCAACTGTATCTTCAGCATCTTTTACTTTTCACACAGCGTTTCTGGGTGAGCAAAAACAGGA	185
3479	Qy	AGGCAAAATGCCGAAAAAAGGGAATAAGGGCGCACACGGAAATGTTGAAATACTCATACTC	3538
184	Db	AGGCAAAATGCCGAAAAAAGGGAATAAGGGCGCACACGGAAATGTTGAAATACTCATACTC	125
3539	Qy	TTCTTTTTCAAATATTATGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA	3598
124	Db	TTCTTTTTCAAATATTATGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA	65
3599	Qy	TTTGAATGTATTAGAAAAATAAACAAATAGGGGTTCCCGCGCACATTTTCCCGAAAAAGTG	3658
64	Db	TTTGAATGTATTAGAAAAATAAACAAATAGGGGTTCCCGCGCACATTTTCCCGAAAAAGTG	5
3659	Qy	CCAC 3662	
4	Db	CCAC 1	

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 13:39:19 ; Search time 610.194 Seconds
(without alignments)
9819.908 Million cell updates/sec

Title: US-10-826-523-41

Perfect score: 3662

Sequence: 1 ctaattgtaagcgttaata.....attccccgaaagtgcac 3662

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2639	72.1	3681	1	US-07-924-028A-6
2	2623.6	71.6	5314	1	US-07-924-028A-4
3	2620.6	71.6	3699	3	US-08-846-538-6
4	2620.6	71.6	3699	3	US-09-503-222-6
5	2469.4	67.4	5277	4	US-09-890-848-16
6	2448.8	66.9	3956	4	US-09-402-266B-21
7	2305.2	62.9	4102	4	US-09-402-266B-20
8	2302	62.9	4088	4	US-09-402-266B-18
9	2279.2	62.2	4119	4	US-09-993-170-1
10	2172.8	59.3	7102	3	US-09-138-024-20
11	2172.8	59.3	7102	3	US-09-404-066-20
12	2172.8	59.3	7102	3	US-09-573-322-20
13	2172.8	59.3	7333	3	US-09-138-024-21
14	2172.8	59.3	7333	3	US-09-404-066-21
15	2172.8	59.3	7333	3	US-09-573-322-21
16	2170.6	59.3	6206	2	US-08-474-169-3
17	2170.6	59.3	6206	2	US-08-474-169-2
18	2170.2	59.3	2973	4	US-09-402-266B-17
19	2170.2	59.3	3789	3	US-09-075-019-8
20	2170.2	59.3	4583	4	US-09-402-266B-19
21	2170	59.3	2961	3	US-08-446-935-6
22	2170	59.3	4145	3	US-08-651-472-62
23	2170	59.3	4145	3	US-08-358-928-62
24	2170	59.3	4277	3	US-08-651-472-63
25	2170	59.3	4277	3	US-08-358-928-63
26	2170	59.3	5356	3	US-08-446-935-1
27	2170	59.3	5353	3	US-08-651-472-72

C 28 2170 59.3 5532 3 US-08-358-928-72 Sequence 72, Appli
C 29 2170 59.3 5534 3 US-08-452-267-3 Sequence 3, Appli
C 30 2170 59.3 5534 3 US-09-123-644-3 Sequence 3, Appli
C 31 2170 59.3 6811 3 US-08-651-472-67 Sequence 67, Appli
C 32 2170 59.3 6811 3 US-08-358-928-67 Sequence 67, Appli
C 33 2170 59.3 6926 3 US-08-651-472-69 Sequence 69, Appli
C 34 2170 59.3 6926 3 US-08-358-928-69 Sequence 69, Appli
C 35 2170 59.3 7090 4 US-09-714-550-18 Sequence 18, Appli
C 36 2170 59.3 7214 4 US-09-438-268-1 Sequence 1, Appli
C 37 2170 59.3 7432 4 US-09-367-891A-6 Sequence 6, Appli
C 38 2170 59.3 8560 4 US-09-936-572-11 Sequence 11, Appli
C 39 2170 59.3 11233 3 US-08-980-832-27 Sequence 27, Appli
C 40 2170 59.3 11233 4 US-09-920-923B-27 Sequence 27, Appli
C 41 2169 59.2 3737 2 US-08-784-208-1 Sequence 1, Appli
C 42 2169 59.2 3796 4 US-09-470-661A-32 Sequence 32, Appli
C 43 2169 59.2 4399 4 US-09-911-781-27 Sequence 27, Appli
C 44 2169 59.2 4399 4 US-10-400-902-27 Sequence 27, Appli
C 45 2169 59.2 6295 2 US-08-659-206A-4 Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-07-924-028A-6

; Sequence 6, Application US/07924028A

; Patent No. 5470573

; GENERAL INFORMATION:

; APPLICANT: Lubitz Werner, Szostak, Michael P.

; TITLE OF INVENTION: CARRIER-BOUND RECOMBINANT PROTEINS, PROCESS

; TITLE OF INVENTION: FOR THE PRODUCTION AND USE AS IMMUNOGENS AND VACCINES

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felfe & Lynch

; STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/924,028A

; FILING DATE: 30-SEP-1992

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/EP91/00308

; FILING DATE: 02-FEB-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: DE 40 05 874

; FILING DATE: 24-FEB-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Hanson, No. 5470573man D.

; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: HUBR 1027

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3681 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-07-924-028A-6

Query Match 72.1%; Score 2639; DB 1; Length 3681;

Best Local Similarity 83.5%; Pred. No. 0;

Matches 3075; Conservative 0; Mismatches 580; Indels 28; Gaps 6;

QY 3 AAATTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCAT 62
Db |||||
QY 1 AAATTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCAT 60
Db |||||
QY 63 TTTTAAACCAATAGCGCGAAATTCGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGA 122
Db |||||
QY 61 TTTTAAACCAATAGCGCGAAATTCGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGA 120
Db |||||
QY 123 TAGGGTTGAGTGTGTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTTGGACTCCA 182
Db |||||
QY 121 TAGGGTTGAGTGTGTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTTGGACTCCA 180
Db |||||
QY 183 ACGTCAAAAGCGGAAACACCGTCTATCAGGGGATGGCCCACTAGCTGAACCAATCACCCCT 242
Db |||||
QY 181 ACGTCAAAAGCGGAAACACCGTCTATCAGGGGATGGCCCACTAGCTGAACCAATCACCCCT 240
Db |||||
QY 243 AATCAAGTTTTTGGGGTTCAGGGTTCGAGTTCGCTTAAGACACTAAATCGGAACCCCTAAAGGGAGCC 302
Db |||||
QY 241 AATCAAGTTTTTGGGGTTCGAGTTCGCTTAAGACACTAAATCGGAACCCCTAAAGGGAGCC 300
Db |||||
QY 303 CCGGATTTAGAGCTTCAGCGGAAAGCGCGGAAAGCTGSCGAGAAAGGAAGGAAGAAG 362
Db |||||
QY 301 CCGGATTTAGAGCTTCAGCGGAAAGCGCGGAAAGCTGSCGAGAAAGGAAGGAAGAAG 360
Db |||||
QY 363 CGAAAGGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCAACGCTGCGCGTAAACCA 422
Db |||||
QY 361 CGAAAGGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCAACGCTGCGCGTAAACCA 420
Db |||||
QY 423 CACCGCGCGCTTAATGCGCGCTACAGGGGCGGTCCCAATTCGCGCATTCAGGCTGCGCA 482
Db |||||
QY 421 CACCGCGCGCTTAATGCGCGCTACAGGGGCGGTCCCAATTCGCGCATTCAGGCTACGCA 480
Db |||||
QY 483 ACTGTTGGGAAGGCGATCGGTGCGGCGCTCTTCGCTATTACGCCAGCTGGCGGAAGGGG 542
Db |||||
QY 481 ACTGTTGGGAAGGCGATCGGTGCGGCGCTCTTCGCTATTACGCCAGCTGGCGGAAGGGG 540
Db |||||
QY 543 GATGCTGTCGAAGGCGATTAAGTTGGGTAAAGCGGAGGTTTTCCAGTCAACGCTGTGTA 602
Db |||||
QY 541 GATGCTGTCGAAGGCGATTAAGTTGGGTAAAGCGGAGGTTTTCCAGTCAACGCTGTGTA 600
Db |||||
QY 603 AAACGAGCGCAGTAGCGCGCTCGTTCAATTCAGCTTTTGAACCCGTTGGAGGAGCGGC 662
Db |||||
QY 601 AAACGAGCGCAGTAGCGCGCTCGTTCAATTCAGCTTTTGAACCCGTTGGAGCTCCACCGCG 660
Db |||||
QY 663 AGACTCGCGTGCATAATGTTTACAGCGTCATCGAGCAGATGAAGTGTGCGACACGC 722
Db |||||
QY 661 GTGGCGCGCGCTCTAGTATGTTGCTCTCAGTACATCTGCTCTGATGCCCATAGTTA 720
Db |||||
QY 723 TGCAGAACACGCGCTAGATTAAACCTTAGAAAGATAATCATATTGTGACGTACGTAAAG 782
Db |||||
QY 721 AGCCAGTATATACCTCCGCTATCGTACGTGACCTGGGTCTGCTGCCGGCATCCGCTTACAGACA 840
Db |||||
QY 843 GGTACCGGCGCCCTCGAGGTCGACGTATCGA- ----TAAGCTTGATATCAATTC 896
Db |||||
QY 841 AGCTGTGACCGTCTCCGGAGCTGATGTGTGAGCAATTCGTCCTTAAGTAAGCAATTGCTG 900
Db |||||
QY 897 CTGACGCGCGGGGATCCACTAGTTCTTAGAGCGGCGCCACCGGTGGAGCTCCAGCTT 956
Db |||||
QY 901 CGCGAGGCGAGTAAGGTCGATGCTTTGTGAGCAATTCGTCCTTAAGTAAGCAATTGCTG 960
Db |||||
QY 957 TTGTTCCCTTTAGTAGGGTTAATTAGATCCCAT- ----CGTCAATTTTACGAGAC 1009
Db |||||
QY 961 TAAAGTCGTCAGTGTGCGGATCACCGCTTCCAGTAGCGACAGAAAGCAATTGATTTGTTAA 1020
Db |||||
QY 1010 TATCTTTCTAGGGTTAACTAGCTGCATCAGGATCATATCTGTCGGGTCTTTTTCGGCT 1069
Db |||||
QY 1021 TTTTCAGAGAAAGATCGCGAGGAAGATCAATACATAAGAGTTGAATCTTTTGTGTCT 1080
Db |||||
QY 1070 CAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGACCGCGCTTTTC 1129
Db |||||

Db |||||
QY 1081 TCGACATGGTAATCTCATGTGTTGAATGGCCCTAGAGGATCCGGCCAAGCTTGCATGCC 1140
Db |||||
QY 1130 CCGCAGGTTGAGCGGCGCATGGAAAGAGTTTCCCGAGGATGACTGCTGCTGATTCAGCT .1189
Db |||||
QY 1141 TCGAGGTGCATCTTAGAGGATCCCGACGCTCGACGCCATTAATAATGTTTCCGTAAT 1200
Db |||||
QY 1190 TCGAGCAAAACGACAGCTTTTACCATGATGATTCGGGAAGGTGGGCATGCACGCGCTTTAA 1249
Db |||||
QY 1201 TCGAGCGCTTTCATGATGAGACAGCGCGCTTT- --GAATGTTGACGGGATGAACATATAAG 1258
Db |||||
QY 1250 CGGTCAACTGTTTCTGAGGCGACCTGGGATACCAAGTTCGTGCGGCTTTTCCGACACA 1309
Db |||||
QY 1259 CAATGACGCGAGCAATAAACTCAACAGG- --AGCAGGAAAGCGAGGTATCCCAAAAGT 1315
Db |||||
QY 1310 GTTCCGGATGTCAGCGGAGCGCATCAGCAACCGGAACAATACCGGCGACAG- ---- 1363
Db |||||
QY 1316 CCAGCGTACCAATAAACGCAAGCCTCAACGACGCGACGACGAGCGGTGAGTAGCAA 1375
Db |||||
QY 1364 CCGGAATCGCGTGGCGGTGTGAGATTAATGACAGCGGTGGCGCTGGGATATTCGT 1423
Db |||||
QY 1376 TCCAAACTTGTACTCTGTCAGAAATCGAATCATCTTCGTTAAATCCAAAACGGCAG 1435
Db |||||
QY 1424 CAGCGAGAACGGGTAT- ----CTGGCTGGATGCCGCAAAATGGACATGGATACCCCGTG 1479
Db |||||
QY 1436 AAGCCTGAATGAGAATTCGACCTCGAGGGGGGCGCGGTACCCAGCTTTGTTCCCTTTA 1495
Db |||||
QY 1480 AGTTACCGCGGCGCGCTTGGCGTAAATCATGTCATAGCTGTTTCTGTCGTAATG 1539
Db |||||
QY 1496 GTGAGGGTTAATTCGAGCTTTGGCGTAAATCATGTCATAGCTGTTTCTGTGTAATG 1555
Db |||||
QY 1540 TTATCCGCTCAAAATTCACACAAACATACGAGCCGGAAGCATAAAGTGTAAAGCTGGGG 1599
Db |||||
QY 1556 TTATCCGCTCAAAATTCACACAAACATAGGAGCCGGAAGCATAAAGTGTAAAGCTGGGG 1615
Db |||||
QY 1600 TGCCTAATAGTGAAGTAACTCAACATTAATTTGCTGTTGCGTCACTGCGCGCTTTCAGT 1659
Db |||||
QY 1616 TGCCTAATAGTGAAGTAACTCAACATTAATTTGCTGTTGCGTCACTGCGCGCTTTCAGT 1675
Db |||||
QY 1660 GGAACACCTGTCGTCAGCTGCATTAATGATCGGCCAAACGCGCGGAGAGCGGTTT 1719
Db |||||
QY 1676 GGAACACCTGTCGTCAGCTGCATTAATGATCGGCCAAACGCGCGGAGAGCGGTTT 1735
Db |||||
QY 1720 GCGTATTCGGCGCTTTCGCTCTCCTCACTCACTCGCTCGCTCGCTCGTTCGGCT 1779
Db |||||
QY 1736 GCGTATTCGGCGCTTTCGCTCTCCTCACTCACTCGCTCGCTCGCTCGTTCGGCT 1795
Db |||||
QY 1780 GCGCGAGCGGTATCAGCTCACTCAAGCGGTAATACGTTTATCCACAGAAATCAGGGA 1839
Db |||||
QY 1796 GCGCGAGCGGTATCAGCTCACTCAAGCGGTAATACGTTTATCCACAGAAATCAGGGA 1855
Db |||||
QY 1840 TAAACGAGAAAGAACATGTGAGCAAAAGCGCAGCAAAAGCGCAGAAACCGTAAAGAGC 1899
Db |||||
QY 1856 TAAACGAGAAAGAACATGTGAGCAAAAGCGCAGCAAAAGCGCAGAAACCGTAAAGAGC 1915
Db |||||
QY 1900 GCGGTTGCTGGCGTTTTTCCATAGGCTCCGCGCTCGACGAGCATCAAAAATTCGACG 1959
Db |||||
QY 1916 GCGGTTGCTGGCGTTTTTCCATAGGCTCGGCGCTCGACGAGCATCAAAAATTCGACG 1975
Db |||||
QY 1960 CTCAGTCAAGGTTGGCGAAACCCGACAGGACTATAAGATACACAGGCTTTCCCGCTGG 2019
Db |||||
QY 1976 CTCAGTCAAGGTTGGCGAAACCCGACAGGACTATAAGATACACAGGCTTTCCCGCTGG 2035
Db |||||
QY 2020 AAGCTCCCTCGTGGCTCTCTGTTCCGACCTCGCTTACCGGATACCTGTCGCGCTT 2079
Db |||||
QY 2036 AAGCTCCCTCGTGGCTCTCTGTTCCGACCTCGCTTACCGGATACCTGTCGCGCTT 2095
Db |||||
QY 2080 TCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGCT 2139
Db |||||
QY 2096 TCTCCCTTCGGAAGCGTGGCGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGCT 2155
Db |||||
QY 2140 GTAGTCTGTTCCGTCAGCTGGGCTGTGTGCAAGAACCCCGCTTTCAGCGCGAGCGCTG 2199
Db |||||

Db 2156 GTAGTGGTTCGCTCCAGCTGGCTGTGTGACGAACCCCGTTTCAGCCGACCGCTG 2215
Qy 2200 CGCTTATCCGGTAACTATCGTCTTGAGTCAACCCCGGTAAAGACACGACTTATCGCACT 2259
Db 2216 CGCTTATCCGGTAACTATCGTCTTGAGTCAACCCCGGTAAAGACACGACTTATCGCACT 2275
Qy 2260 GGCAGGACCACTGGTAAACAGGATTAAGCAGGAGGATAGTAGGCGGTCTACAGAGTT 2319
Db 2276 GGCAGGACCACTGGTAAACAGGATTAAGCAGGAGGATAGTAGGCGGTCTACAGAGTT 2335
Qy 2320 CTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGGCTCT 2379
Db 2336 CTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGGCTCT 2395
Qy 2380 GCTGAAGCCAGTTACCTTCGGAAAAAGAGTGTGTAGCTCTTGTGATTCGGGCAAAAACAAACAC 2439
Db 2396 GCTGAAGCCAGTTACCTTCGGAAAAAGAGTGTGTAGCTCTTGTGATTCGGGCAAAAACAAACAC 2455
Qy 2440 CGCTGGTAGCGGTGGTCTTCTGTTGCAAGCAGCAGATTAACGGCGAGAAAAAAGATC 2499
Db 2456 CGCTGGTAGCGGTGGTCTTCTGTTGCAAGCAGCAGATTAACGGCGAGAAAAAAGATC 2515
Qy 2500 TCAAGAGATCCTTTGATCTTTCTACGGGTCTGACGCTCAGTGGAAACGAAACTCACG 2559
Db 2516 TCAAGAGATCCTTTGATCTTTCTACGGGTCTGACGCTCAGTGGAAACGAAACTCACG 2575
Qy 2560 TTAAGGATTTGGTCAATGAGATTAATCAAAAGGATCTTCACCTAGATCTTTTAAATTA 2619
Db 2576 TTAAGGATTTGGTCAATGAGATTAATCAAAAGGATCTTCACCTAGATCTTTTAAATTA 2635
Qy 2620 AAAATGAAGTTTAAATCAATCTAAAGTATATATAGTAAACTTGGTCTGACAGTTACCA 2679
Db 2636 AAAATGAAGTTTAAATCAATCTAAAGTATATATAGTAAACTTGGTCTGACAGTTACCA 2695
Qy 2680 ATCTTAAATCAGTAGGACCTATCTCAGGATCTGTCTATTTTCTGTTTTCATCCATAGTTGC 2739
Db 2696 ATCTTAAATCAGTAGGACCTATCTCAGGATCTGTCTATTTTCTGTTTTCATCCATAGTTGC 2755
Qy 2740 CTGACTCCCCGTCGTGTAGATACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGC 2799
Db 2756 CTGACTGCCGTCGTGTAGATACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGC 2815
Qy 2800 TGAATGATACCGGAGACCCAGCTCAGCGGTCCAGATTTATCAGCAATAAACGAGCC 2859
Db 2816 TGAATGATACCGGAGACCCAGCTCAGCGGTCCAGATTTATCAGCAATAAACGAGCC 2875
Qy 2860 AGCCGAAGGCGGAGCGAGAGTGTCTGCAACTTTATCCGCTCCATCCAGTCTAT 2919
Db 2876 AGCCGAAGGCGGAGCGAGAGTGTCTGCAACTTTATCCGCTCCATCCAGTCTAT 2935
Qy 2920 TAATTGTTGCCGGAAGCTAGAGTAAGTAGTTGCCAGTTAATAGTTTGGCAACGTTGT 2979
Db 2936 TAATTGTTGCCGGAAGCTAGAGTAAGTAGTTGCCAGTTAATAGTTTGGCAACGTTGT 2995
Qy 2980 TGCATTTGCTACAGGCATCGTGTGTCA CGCTCGTCTGTTTGGTATGGCTTCATTCAGTCT 3039
Db 2996 TGCATTTGCTACAGGCATCGTGTGTCA CGCTCGTCTGTTTGGTATGGCTTCATTCAGTCT 3055
Qy 3040 CGCTTCCACAGTCAAGGCGAGTACATCATCCCCCATGTTGTGCAAAAAGCGGTAG 3099
Db 3056 CGCTTCCACAGTCAAGGCGAGTACATCATCCCCCATGTTGTGCAAAAAGCGGTAG 3115
Qy 3100 CTCTTTCGGTCTCCGATCGTTGTGCAAGTAAAGTTGGCCGAGTGTATCACTCATGTT 3159
Db 3116 CTCTTTCGGTCTCCGATCGTTGTGCAAGTAAAGTTGGCCGAGTGTATCACTCATGTT 3175
Qy 3160 TATGGCAGCATCGATAAATCTCTTACTGTCTATGTCGCATCCGTAAGATGCTTTTCTGTGAC 3219
Db 3176 TATGGCAGCATCGATAAATCTCTTACTGTCTATGTCGCATCCGTAAGATGCTTTTCTGTGAC 3235
Qy 3220 TGGTGTACTCAACCAAGTCACTTCTGAGATAGTGTATCGGGGACCGAGTTGCTCTTG 3279
Db 3236 TGGTGTACTCAACCAAGTCACTTCTGAGATAGTGTATCGGGGACCGAGTTGCTCTTG 3295

Qy 3280 CCGGGCTCAATACGGGATTAATACCGGCCACATAGCAGAACTTTAAAAGTGTCTCATCAT 3339
Db 3296 CCGGGCTCAATACGGGATTAATACCGGCCACATAGCAGAACTTTAAAAGTGTCTCATCAT 3355
Qy 3340 TGCAGAAACGTTCTTTCCGGGGGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTC 3399
Db 3356 TGCAGAAACGTTCTTTCCGGGGGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTC 3415
Qy 3400 GATGTAACCCACTCGTGACCCCAACTGATCTTCAGCATCTTTTACCTTTCCACGCGTTTC 3459
Db 3416 GATGTAACCCACTCGTGACCCCAACTGATCTTCAGCATCTTTTACCTTTCCACGCGTTTC 3475
Qy 3460 TGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGAAATAGGGCGGACACCGAA 3519
Db 3476 TGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGAAATAGGGCGGACACCGAA 3535
Qy 3520 ATGTTGAATACTCATACTCTCTCTTTTCAATATTTTGAAGCATTTTATCAGGGTTATTG 3579
Db 3536 ATGTTGAATACTCATACTCTCTCTTTTCAATATTTTGAAGCATTTTATCAGGGTTATTG 3595
Qy 3580 TCTCATGAGCGGATACATATTTGAATGTTTATAGAAAAATAAATAGGGGTTCCGCG 3639
Db 3596 TCTCATGAGCGGATACATATTTGAATGTTTATAGAAAAATAAATAGGGGTTCCGCG 3655
Qy 3640 CACATTTCCCGAAAAAGTGCCAC 3662
Db 3656 CACATTTCCCGAAAAAGTGCCAC 3678

RESULT 2

US-07-924-028A-4
; Sequence 4, Application US/07924028A
; Patent No. 5470573
; GENERAL INFORMATION:
; APPLICANT: Lubitz Werner, Szostak, Michael P.
; TITLE OF INVENTION: CARRIER-BOUND RECOMBINANT PROTEINS, PROCESS
; TITLE OF INVENTION: FOR THE PRODUCTION AND USE AS IMMUNOGENS AND VACCINES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/924,028A
; FILING DATE: 30-SEP-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP91/00308
; FILING DATE: 02-FEB-1991
; PRIOR APPLICATION DATA: DE 40 05 874
; APPLICATION NUMBER: 24-FEB-1990
; FILING DATE: 24-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5470573man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5314 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear			
US-07-924-028A-4			
Query Match		71.6%; Score 2623.6; DB 1; Length 5314;	
Best Local Similarity		83.4%; Pred. No. 0;	
Matches 3059; Conservative		0; Mismatches 579; Indels 28; Gaps 6;	
Qy	20	ATTTTGTAAATTCGCGTAAATTTTCTTAAATCAGCTCAATTTTAAACCAATAGGCC	79
Db	1651	ATTTTGTAAATTCGCGTAAATTTTGTAAATCAGCTCAATTTTAAACCAATAGGCC	1710
Qy	80	GAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGATAGGTTGAGTGT	139
Db	1711	GAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGATAGGTTGAGTGT	1770
Qy	140	CCAGTTTGGAACAAGATCCACTATTAAGAACGTGGACTCCAAGTCAAAAGGCGMAAA	199
Db	1771	CCAGTTTGGAACAAGATCCACTATTAAGAACGTGGACTCCAAGTCAAAAGGCGMAAA	1830
Qy	200	ACCGTCTATCAGGGGATGGCCACTACGTGAACCATCACCCCTAATCAAGTTTTTGGG	259
Db	1831	ACCGTCTATCAGGGGATGGCCACTACGTGAACCATCACCCCTAATCAAGTTTTTGGG	1890
Qy	260	TCAGGTGCGGTAAAGCACTAAATCGGAACCTTAAAGGAGCCCCCGATTTAGAGCTTGA	319
Db	1891	TCAGGTGCGGTAAAGCACTAAATCGGAACCTTAAAGGAGCCCCCGATTTAGAGCTTGA	1950
Qy	320	CGGGAAAGCCCGGCAACGTGCGGAGAAAGGAGGAGAAACGAAAGGAGCGGCGCT	379
Db	1951	CGGGAAAGCCCGGCAACGTGCGGAGAAAGGAGGAGGAGAAACGAAAGGAGCGGCGCT	2010
Qy	380	AGGGCGCTGGCAAGTGTAGCGTACGCTGCGCTGAACCAACACACACCCCGCGCTTAAT	439
Db	2011	AGGGCGCTGGCAAGTGTAGCGTACGCTGCGCTGAACCAACACACCCCGCGCTTAAT	2070
Qy	440	GCGCGCTACAGGGCGGCTCCATTGCGCATTCAGCTGCGCAACTGTGTGGAAAGGCGA	499
Db	2071	GCGCGCTACAGGGCGGCTCCATTGCGCATTCAGCTGCGCAACTGTGTGGAAAGGCGA	2130
Qy	500	TCGGTGGCGGCTCTTCGCTATTAGCCAGCTGCGGAAGGGGATGTCTGCAAGGCGA	559
Db	2131	TCGGTGGCGGCTCTTCGCTATTAGCCAGCTGCGGAAGGGGATGTCTGCAAGGCGA	2190
Qy	560	TTAAGTTGGTTAACCGCAGGGTTTTCCAGTCACGACAGTTGTAAACGACGCGCCAGTGAG	619
Db	2191	TTAAGTTGGTTAACCGCAGGGTTTTCCAGTCACGACAGTTGTAAACGACGCGCCAGTGAA	2250
Qy	620	CAGCGCTCGTTCAATTCAGCTTTTGAACCCGTGGAGGACGGGACAGCTCGCGGTGCAAT	679
Db	2251	TTGTAAACGACTCACTATAGGGCGAAATTGGAGCTCCACCGCGTGGCGCGCTCTAGT	2310
Qy	680	GTGTTTACAGCGTGTAGGAGCAGATGAAGATGCTCGACACCGCTCGAGAACACGAGCTA	739
Db	2311	ATGGTGCACTCTCAGTACAAATCTGCTGTGATGCCGATAGTTAAGCCAGTATATACACTC	2370
Qy	740	GATTAACCCCTGAAAGATAATCATATTGTGACGTACGTTTAAAGATAATCATCGCTAAAT	799
Db	2371	CGCTATCGTACTGACTGGTCAATGGCTGCGCCCGGACACCGGCCAACCCCGCTGAGC	2430
Qy	800	TGACGCATGGGATCTGTAAATACGACTCACTATAGGGCGAATTTGGGTACCGGGCCCCCCT	859
Db	2431	CGCCCTGACGGGCTGTCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGCTCTCG	2490
Qy	860	CGAGTTCGAGGTATCGA-----TAAGCTTGATATCGAATTCCTGCAGCCCGGGGATC	913
Db	2491	GGAGTGCATGTGTGAGAGGTTTTTACCGCTCATCACCCGAAACGCGGAGGAGTAAAGTTC	2550
Qy	914	CACCTAGTTCTAGAGCGGCGCCACCGCGTGGAGCTCCAGCTTTTGTTCCTTTTGTAGTGA	973
Db	2551	GGATGCTTTGTGAGCAATTCGTCCTTAAAGTAAGCAATTTGCTGTAAGTCGCTCACTGTGC	2610
Qy	974	GGTAAATAGATCCCAT-----GGTCAATTTTACGCAGACTATCTTTCTAGGGTTAA	1026

QY 2097 TGCGGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTTCGGTGTAGGTCTGTTTCGTTCCCA 2156
DB 3746 TGCGGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTTCGGTGTAGGTCTGTTTCGTTCCCA 3805
QY 2157 AGCTGGCTGTGTGACGACGACCCCGTTTCAGCCGACCGCTGCGCTTATCGGTTAACT 2216
DB 3806 AGCTGGCTGTGTGACGACGACCCCGTTTCAGCCGACCGCTGCGCTTATCGGTTAACT 3865
QY 2217 ATCGCTTTGAGTCCAAACCGGTAAGACAGCACTTATCGCACCTGGCAGCAGCCACTGGTA 2276
DB 3866 ATCGCTTTGAGTCCAAACCGGTAAGACAGCACTTATCGCACCTGGCAGCAGCCACTGGTA 3925
QY 2277 ACAGGATAGCAGAGGAGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGTGTGCGCTTA 2336
DB 3926 ACAGGATAGCAGAGGAGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGTGTGCGCTTA 3985
QY 2337 ACTACGGCTACACTAGAAGCAGATATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCT 2396
DB 3986 ACTACGGCTACACTAGAAGCAGATATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCT 4045
QY 2397 TCGGAAAAAGAGTGTGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGTGGTT 2456
DB 4046 TCGGAAAAAGAGTGTGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGTGGTT 4105
QY 2457 TTTTGTGTGCAAGCAGCAGATTACCGCGCAGAAAAAAGATCTCAAGAGATCTCTTTGA 2516
DB 4106 TTTTGTGTGCAAGCAGCAGATTACCGCGCAGAAAAAAGATCTCAAGAGATCTCTTTGA 4165
QY 2517 TCTTTTCTACGGGCTGACGCTCAGTGGAAAGAACTCAGTTAAGGATTTTGGTCA 2576
DB 4166 TCTTTTCTACGGGCTGACGCTCAGTGGAAAGAACTCAGTTAAGGATTTTGGTCA 4225
QY 2577 TGAGATTATCAAAAAGGATCTTCACCTAGATCTCTTTTAAATTAATAAGATTTTAAAT 2636
DB 4226 TGAGATTATCAAAAAGGATCTTCACCTAGATCTCTTTTAAATTAATAAGATTTTAAAT 4285
QY 2637 CAATCTAAAGTATATAGTAACTTGTGTGACAGTTACCAATGCTTAAATCAGTGAGG 2696
DB 4286 CAATCTAAAGTATATAGTAACTTGTGTGACAGTTACCAATGCTTAAATCAGTGAGG 4345
QY 2697 CACCTATCTCAGGATCTGTATTTTGTTCATCTAGTTGCTGACTCCCGTCTGT 2756
DB 4346 CACCTATCTCAGGATCTGTATTTTGTTCATCTAGTTGCTGACTCCCGTCTGT 4405
QY 2757 AGATACTACGATACGGGAGGCTTACCATCTGGCCCGAGTGTGCAATGATACCCGAG 2816
DB 4406 AGATACTACGATACGGGAGGCTTACCATCTGGCCCGAGTGTGCAATGATACCCGAG 4465
QY 2817 ACCACGCTCACGGCTCCAGATTTATCAGCAATAAACCCAGCAGCCGGAAGGCGGAGC 2876
DB 4466 ACCACGCTCACGGCTCCAGATTTATCAGCAATAAACCCAGCAGCCGGAAGGCGGAGC 4525
QY 2877 GCAGAGTGTCTGCAACTTATCCGCTCCATCCAGTCTATTATTTGTCGGGAG 2936
DB 4526 GCAGAGTGTCTGCAACTTATCCGCTCCATCCAGTCTATTATTTGTCGGGAG 4585
QY 2937 CTAGAGTAAAGTGTCCGAGTTAATAGTTTGGCAACGTTGTGCAATGCTTACAGGCA 2996
DB 4586 CTAGAGTAAAGTGTCCGAGTTAATAGTTTGGCAACGTTGTGCAATGCTTACAGGCA 4645
QY 2997 TCGTGGTGTACGCTCGTGTGTGTGATGCTTTCATTCAGCTCCGGTTCCTCAACGATCAA 3056
DB 4646 TCGTGGTGTACGCTCGTGTGTGTGATGCTTTCATTCAGCTCCGGTTCCTCAACGATCAA 4705
QY 3057 GGGAGTTACATGATCCCGATTTGTGCAAAAAGGGTGTAGTCTCTTCGGTCTCCGA 3116
DB 4706 GGGAGTTACATGATCCCGATTTGTGCAAAAAGGGTGTAGTCTCTTCGGTCTCCGA 4765
QY 3117 TCGTGTGCAAGTAAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATA 3176
DB 4766 TCGTGTGCAAGTAAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATA 4825

QY 3177 ATTCTTCTACTCTCATGCCATCGTAAGATGCTTTTCTGACTGCTGACTCAACCA 3236
DB 4826 ATTCTTCTACTCTCATGCCATCGTAAGATGCTTTTCTGACTGCTGACTCAACCA 4885
QY 3237 AGTCATTTCTGAGAAAGTAGTGTATGCGCGCACCGAGTTGCTCTTCCCGCGGTCAATACGGG 3296
DB 4886 AGTCATTTCTGAGAAAGTAGTGTATGCGCGCACCGAGTTGCTCTTCCCGCGGTCAATACGGG 4945
QY 3297 ATAATAACCGCGCACATAGCAGAACTTTTAAAGTGTCTCATTTGGAACCGTTCTTCGG 3356
DB 4946 ATAATAACCGCGCACATAGCAGAACTTTTAAAGTGTCTCATTTGGAACCGTTCTTCGG 5005
QY 3357 GCGGAAAACCTCTCAAGGATCTTACCGCTCTTGAGATCCAGTTTCGATGTAACCCACTCGTG 3416
DB 5006 GCGGAAAACCTCTCAAGGATCTTACCGCTCTTGAGATCCAGTTTCGATGTAACCCACTCGTG 5065
QY 3417 CACCCAACTGATCTTTCAGCATCTTTTACCTTTTACCCAGCGTTTCTGGGTGAGCAAAACAG 3476
DB 5066 CACCCAACTGATCTTTCAGCATCTTTTACCTTTTACCCAGCGTTTCTGGGTGAGCAAAACAG 5125
QY 3477 GAAGCAAAATGCGCAAAAGGGAATTAAGGGCGACAGGAAATGTTGAATCTCATAC 3536
DB 5126 GAAGCAAAATGCGCAAAAGGGAATTAAGGGCGACAGGAAATGTTGAATCTCATAC 5185
QY 3537 TCTTCTCTTTTCAATATTTTGAAGCATTTTATCAGGTTTATTCAGGTTTATTCATGAGCGATACA 3596
DB 5186 TCTTCTCTTTTCAATATTTTGAAGCATTTTATCAGGTTTATTCAGGTTTATTCATGAGCGATACA 5245
QY 3597 TATTTGAATGTTTATAGAAAAATAAATAAGGGTTCCGGCGACATTTCCCGGAAAG 3656
DB 5246 TATTTGAATGTTTATAGAAAAATAAATAAGGGTTCCGGCGACATTTCCCGGAAAG 5305
QY 3657 TGCCAC 3662
DB 5306 TGCCAC 5311

RESULT 3
US-08-646-538-6
; Sequence 6, Application US/08646538
; Patent No. 6027881
; GENERAL INFORMATION:
; APPLICANT: Pavlakis, George N.
; APPLICANT: Gaitanaris, George A.
; APPLICANT: Stauber, Roland H.
; APPLICANT: Vournakis, John N.
; TITLE OF INVENTION: Mutant Aequorea victoria Fluorescent
; TITLE OF INVENTION: Proteins Having Increased Cellular Fluorescence
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,538
; FILING DATE: No. 6027881 yet assigned
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 015280-249000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

; LENGTH: 3699 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA

; FEATURE:

; NAME/KEY: -

; LOCATION: 1..3699

; OTHER INFORMATION: /note= "pBSGFP"

US-08-646-538-6

Query Match

Best Local Similarity 71.6%; Score 2620.6; DB 3; Length 3699;

Matches 3078; Conservative 0; Mismatches 579; Indels 40; Gaps 7;

QY 3 AAATTGTAAGCGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTCAT 62
DB 3 AAATTGTAACGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTCAT 62
QY 63 TTTTAAACCAATAGCGGAAATCGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGA 122
DB 63 TTTTAAACCAATAGCGGAAATCGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGA 122
QY 123 TAGGGTTGAGTGTGTTCAGTTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTCCA 182
DB 123 TAGGGTTGAGTGTGTTCAGTTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTCCA 182
QY 183 AGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTAGCTGAACCATCACCTT 242
DB 183 AGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTAGCTGAACCATCACCTT 242
QY 243 AATCAAGTTTTTGGGTGCGAGTCCGTPAAGCACTAAATCGGAACCCCTAAAGGGAGCC 302
DB 243 AATCAAGTTTTTGGGTGCGAGTCCGTPAAGCACTAAATCGGAACCCCTAAAGGGAGCC 302
QY 303 CCCGATTTAGAGCTTTGACGGGAAACCGCGGAAACGTCGCGAAGAAAGGAAAGAAAG 362
DB 303 CCCGATTTAGAGCTTTGACGGGAAACCGCGGAAACGTCGCGAAGAAAGGAAAGAAAG 362
QY 363 CGAAAGGAGCGGCGTAGGGCGCTGGCAGTGTAGCGGTACGCTGGCGGTAAACCAACA 422
DB 363 CGAAAGGAGCGGCGTAGGGCGCTGGCAGTGTAGCGGTACGCTGGCGGTAAACCAACA 422
QY 423 CACCCGCGCGCTTAATGCGCGCTACAGGGCGCTC --- CCATTGCGCATTCAGCGTGC 479
DB 423 CACCCGCGCGCTTAATGCGCGCTACAGGGCGCTC --- CCATTGCGCATTCAGCGTGC 482
QY 480 GCAACTGTTGGGAGGCGGATCGGTGCGGGCCCTCTTCGCTATTACGCCAGCTGGCGAAAG 539
DB 483 GCAACTGTTGGGAGGCGGATCGGTGCGGGCCCTCTTCGCTATTACGCCAGCTGGCGAAAG 542
QY 540 GGGGATGTCTCAAGGCGATTAAGTTGGGTAAACGCGAGGGTTTTCCAGTCACGACGTT 599
DB 543 GGGGATGTCTCAAGGCGGATTAAGTTGGGTAAACGCGAGGGTTTTCCAGTCACGACGTT 602
QY 600 GTAAACGACGCGCAGTGAAGCGCGCTCGTTTCATTTCAGCTTTTGAACCCGCTGGAGGAGC 659
DB 603 GTAAACGACGCGCAGTGAAGTTGTAAT --- ACGACTCACTATAGGGCGAATTGGGTACCGG 660
QY 660 GGCAGACTCGCGGTGCAATGTGTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACA 719
DB 661 GCCCCCCCTCGAGGTGCGGATCGATGAAGCTTCGATGATCCTTATTGTATAGTTTCATC 720
QY 720 CGCTGCA --- GAACACGACGTAGATTAACCCCTAGAAAGATAATCATATTGTGACGTACG 776
DB 721 CATGCCATGTGTAAATCCAGCAGCTGTTACAAACTCAAGAAGGACCATGTGCTCTCTT 780
QY 777 TTAAGATAATCATCGTAAATATTGACCGATGGGATCTGTAATACGACTCACTATAGGCG 836
DB 781 TTCGTTGGGATCTTTTCGAAAGGGCAGATTGTGTGACAGGTA - ATGGTTGTCTGGTAA 839
QY 837 GAAATTGGGTACCGGGCCCCCCTCGAGGTGCGAGTTCGATTCGATTAAGCTTTGATATCGAATTC 896

DB 840 GGACAGGCGCATCGCCAATTGGAGTATTTGTGTGATAATGGTCTGCTAGTTGAAAGCTTC 899
QY 897 CTGACAGCGCGGGGATCACTAGTT --- CTAGACGCGCGCCACCGCGGTGGAGCTCCA 952
DB 900 CATCTTCAATGTGTCTAAATTTTGAAGTTAACTTTGATTCCTTTTGTGTCTG 959
QY 953 GCTTTTGTCCCTTTAGTGAGGGTTAAJTAGATCCATCGCGTCAATTTTACCGACAGTAT 1012
DB 960 CCATGATGTATACATTGTGTGAGTTATAGTTGTATTCCAATTTGTGTCCAAGAAATGTTTC 1019
QY 1013 CTTTCTAGGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGCGCTCAG 1072
DB 1020 CATCTTCTTTAAATCAATACCTTTTAACTGATTTTAAACAAGGGTATCACCTTCAA 1079
QY 1073 TCATCGCCCAAGCTCGCGCTATCTGGGCATCGGGAGGAAGNAGCCGCTGCTTTCCG 1132
DB 1080 ACTTGACTTCAGCACGTGTCTGTAGTTCCCGTCACTTTTGAAGAAATATAGTTCCTTCT 1139
QY 1133 CGAGGTTCGAAGCGGCATGGAAGAGTTTTCGCGAGTACGCTGCTGCTGCAATTGAGTTGA 1192
DB 1140 GTACATAACCTTCGGGCATGGCACTCTTTGAAAAAGTCAATCGCGTTTCATATGATCGGGT 1199
QY 1193 GCGAAACGCAAGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCAAGCGCTTTAAACGG 1252
DB 1200 ATCTTTGAAAAGCATTTGAACACCCATAAGAGAAAGTAGTGACAAGTGTTCGCCATGGAACAG 1259
QY 1253 TGAACGTGTTCGTTCAGGCCACCTGGGATACCAAGTTTCGTCGGGCTTTTCCGACACAGTT 1312
DB 1260 GTAGTTTTCAGAGTAGTGCAAAATAAATTTAAGGGTAAAGTTTCCGATATGTGTGATCACCTT 1319
QY 1313 CGGATGTGTCAGCCCGAAGCGCATCAGCAACCCGAAACAATACCGCGCAGCAGCCGGAACGTG 1372
DB 1320 CACCTCTCCACTGACAGAAATTTGTGCCCATTAACATCACCATCTAATTTCAACAAGAA 1379
QY 1373 CGGTGCC --- GGTGTGCAGATTAATGACAGGGTGTGGCGCTGGGATAT 1418
DB 1380 TTGGGACAACCTCAGTGAAGAGTTCTTCTCTTTCTAGCCATTTCTTTCGCGCATCGAAT 1439
QY 1419 TACGTACAGGAGGAGCGGTATCCTGGCTGGATGCGCGAGAAATGACATGGATACCCGT 1478
DB 1440 TCCTCAGCCCGGGGATCCACTAGTTCTAGAGCGCGCCCAACCCGCGTGGAGTCCACG 1499
QY 1479 GAGTTACCCCGCGG --- CGCGCTTTGGCGTAACTCATGTGCTAGTCTGTTT 1525
DB 1500 TTTTGTCCCTTAGTGAGGGTTAATTCGAGCTTGGCGTAACTCATGTCTAGTCTGTT 1559
QY 1526 CTTGTGAAATTTGTATTCGCTCACAATTCACAAACATACGAGCGCGGAAGCATAAAG 1585
DB 1560 CTTGTGAAATTTGTATTCGCTCACAATTCACAAACATACGAGCGCGGAAGCATAAAG 1619
QY 1586 TGTAAAGCTGGGTGCTTAATGAGTGAAGTAACTCACAATTAATTTGCGTTCGCTCACCTG 1645
DB 1620 TGTAAAGCTGGGTGCTTAATGAGTGAAGTAACTCACAATTAATTTGCGTTCGCTCACCTG 1679
QY 1646 CCCGCTTTCCAGTCGGGAAACCTGTGTCGCCAGCTGCAATTAATGAATCGGCCCAACGCGCG 1705
DB 1680 CCCGCTTTCCAGTCGGGAAACCTGTGTCGCCAGCTGCAATTAATGAATCGGCCCAACGCGCG 1739
QY 1706 GGGAGAGCGGTTTTCGCTATTTGGGCGCTTTCGCGCTTCTTCGCTCACTGCTGCTGCGC 1765
DB 1740 GGGAGAGCGGTTTTCGCTATTTGGGCGCTTTCGCGCTTCTTCGCTCACTGCTGCTGCGC 1799
QY 1766 TCGGTTCGTTTCGCGTTCGGGAGCGGTATCAGCTCACTCAAGGCGGTAAATACGGTTATCC 1825
DB 1800 TCGGTTCGTTTCGCGTTCGGGAGCGGTATCAGCTCACTCAAGGCGGTAAATACGGTTATCC 1859
QY 1826 ACAGAAATCAGGGATTAACGCGAGGAAAGAAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGG 1885
DB 1860 ACAGAAATCAGGGATTAACGCGAGGAAAGAAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGG 1919
QY 1886 AACCGTAAAGAGCGCGCTTTCGCTGCGGTTTTCATAGGCTCCGCGCCCTCGACGAGCAT 1945

Db 1920 AACGTAAGGCGCGTGTCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCAT 1979
Qy 1946 CACAAAAATCGACGCTCAAGTCAGAGGTGCGAAACCCGACGAGACTATAAAGATACCAG 2005
Db 1980 CACAAAAATCGACGCTCAAGTCAGAGGTGCGAAACCCGACGAGACTATAAAGATACCAG 2039
Qy 2006 GCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCGCTGCGCTTACCGGA 2065
Db 2040 GCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCGCTGCGCTTACCGGA 2099
Qy 2066 TACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGG 2125
Db 2100 TACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGG 2159
Qy 2126 TATCTCAGTTCGCTGAGTGTGCTTCCGCTCAAGCTGGCTGTGTGCAACGAAACCCCGCTT 2185
Db 2160 TATCTCAGTTCGCTGAGTGTGCTTCCGCTCAAGCTGGCTGTGTGCAACGAAACCCCGCTT 2219
Qy 2186 CAGCCCGACCGCTGCGCTTATCCGCTAACTATCGTCTTGAAGTCCAAACCGGTAAAGCAC 2245
Db 2220 CAGCCCGACCGCTGCGCTTATCCGCTAACTATCGTCTTGAAGTCCAAACCGGTAAAGCAC 2279
Qy 2246 GACTTATCGCACTGGCAGCAGCACTGGTAAACAGGATTAGCAGAGCGAGGTATGAGGC 2305
Db 2280 GACTTATCGCACTGGCAGCAGCACTGGTAAACAGGATTAGCAGAGCGAGGTATGAGGC 2339
Qy 2306 GGTGTACAGAGTTCCTGAAGTGGTGGCTTAACCTACGGCTACACTAGAAGGACAGTATTT 2365
Db 2340 GGTGTACAGAGTTCCTGAAGTGGTGGCTTAACCTACGGCTACACTAGAAGGACAGTATTT 2399
Qy 2366 GGTATCTGGCTCTGTGAAGCAGTACCTTCGGAAGAGAGTGGTAGCTCTTGATCC 2425
Db 2400 GGTATCTGGCTCTGTGAAGCAGTACCTTCGGAAGAGAGTGGTAGCTCTTGATCC 2459
Qy 2426 GGCACAAACACCGCTGTAGCGTGGTTTTTTTGTGTTGCAAGCAGCAGATTTAGCGC 2485
Db 2460 GGCACAAACACCGCTGTAGCGTGGTTTTTTTGTGTTGCAAGCAGCAGATTTAGCGC 2519
Qy 2486 AGAAAAAAGGATCTCAAGAAGATCTTTGATCTTTTCAACGGGTCTGACGCTCAGTGG 2545
Db 2520 AGAAAAAAGGATCTCAAGAAGATCTTTGATCTTTTCAACGGGTCTGACGCTCAGTGG 2579
Qy 2546 AACGAAAACTCAGTTAAGGGATTTGGTCATGAGATTATCAAAAAGGATCTTACCTAG 2605
Db 2580 AACGAAAACTCAGTTAAGGGATTTGGTCATGAGATTATCAAAAAGGATCTTACCTAG 2639
Qy 2606 ATCTTTTAAATTAATAAAGTGTAAATCAATCTAAAGTATATATGAGTAAACTTGG 2665
Db 2640 ATCTTTTAAATTAATAAAGTGTAAATCAATCTAAAGTATATATGAGTAAACTTGG 2699
Qy 2666 TCTGACAGTTACCAATGCTTAATCAGTGAAGGCACTATCTCAGCGATCTGTCTATTTCGT 2725
Db 2700 TCTGACAGTTACCAATGCTTAATCAGTGAAGGCACTATCTCAGCGATCTGTCTATTTCGT 2759
Qy 2726 TCATCATAGTTGCTGATCTCCCGCTGTGTAGATACTACGATACGGAGGGCTTACCA 2785
Db 2760 TCATCATAGTTGCTGATCTCCCGCTGTGTAGATACTACGATACGGAGGGCTTACCA 2819
Qy 2786 TCTGGCCCCAGTGTGCAATGATACCGCAGACCCACGCTCACCGGCTCCAGATTATCA 2845
Db 2820 TCTGGCCCCAGTGTGCAATGATACCGCAGACCCACGCTCACCGGCTCCAGATTATCA 2879
Qy 2846 GCAATAAACCCAGCAGCGGAGGGCGGAGCGCAGAGTGGTCTCTGCAACTTTATCCGCC 2905
Db 2880 GCAATAAACCCAGCAGCGGAGGGCGGAGCGCAGAGTGGTCTCTGCAACTTTATCCGCC 2939
Qy 2906 TCATCCAGTCTAATTAATTTGCTCCGGGAAGCTAGAGTAGTGTGCGCAGTTAATAGT 2965
Db 2940 TCATCCAGTCTAATTAATTTGCTCCGGGAAGCTAGAGTAGTGTGCGCAGTTAATAGT 2999
Qy 2966 TTGCGCAAGTGTGCTGCAATGATACCGCAGCAGTGTGCTCAGCTCGCTGTTGGTATG 3025
Db 3000 TTGCGCAAGTGTGCTGCAATGATACCGCAGCAGTGTGCTCAGCTCGCTGTTGGTATG 3059

Qy 3026 GCTTCATTGAGTCCCGTTCCTCCAAACGATCAAGCGAGTTTACATGATCCCCCATGTTGTGC 3085
Db 3060 GCTTCATTGAGTCCCGTTCCTCCAAACGATCAAGCGAGTTTACATGATCCCCCATGTTGTGC 3119
Qy 3086 AAAAAAGCGTTAGCTTCCTTCGGTCTCCGATCGTTGTCAAGAGTAGTGTGGCGCAGTG 3145
Db 3120 AAAAAAGCGTTAGCTTCCTTCGGTCTCCGATCGTTGTCAAGAGTAGTGTGGCGCAGTG 3179
Qy 3146 TTATCACTCATGTTATGCGAGCACTGCATAAATCTCTTACTGTCAATGCCATCCGTAAGA 3205
Db 3180 TTATCACTCATGTTATGCGAGCACTGCATAAATCTCTTACTGTCAATGCCATCCGTAAGA 3239
Qy 3206 TGCTTTTCTGTGACTGGTGAAGTACTCAACCAAGTCAATCTCGAAGATAGTGTATGCGCGCA 3265
Db 3240 TGCTTTTCTGTGACTGGTGAAGTACTCAACCAAGTCAATCTCGAAGATAGTGTATGCGCGCA 3299
Qy 3266 CCGAGTTGCTCTTCGCGCGGCTCAATACGGGATAAATACCGGCCACATAGCAGAACTTTA 3325
Db 3300 CCGAGTTGCTCTTCGCGCGGCTCAATACGGGATAAATACCGGCCACATAGCAGAACTTTA 3359
Qy 3326 AAAGTCTCATCATTTGAAAAACGTTCTTCGGGCGGAAAACTCTCAAGGATCTTACCCTG 3385
Db 3360 AAAGTCTCATCATTTGAAAAACGTTCTTCGGGCGGAAAACTCTCAAGGATCTTACCCTG 3419
Qy 3386 TTGAGATCCAGTTTCGATGTAACCCACTCGTGCAACCAACTGTATCTTCAGCATCTTTTACT 3445
Db 3420 TTGAGATCCAGTTTCGATGTAACCCACTCGTGCAACCAACTGTATCTTTCAGCATCTTTTACT 3479
Qy 3446 TTCACCAGCGTTTCGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATA 3505
Db 3480 TTCACCAGCGTTTCGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATA 3539
Qy 3506 AGGGCGACACGGAATGTTGAATACATCATCTCTCTTTTCAATATTTATGAAGCATT 3565
Db 3540 AGGGCGACACGGAATGTTGAATACATCATCTCTCTTTTCAATATTTATGAAGCATT 3599
Qy 3566 TATCAGGTTATGTTCTCATGAGCGGATACATATTTGAATGTTATGAAAAATAAACAA 3625
Db 3600 TATCAGGTTATGTTCTCATGAGCGGATACATATTTGAATGTTATGAAAAATAAACAA 3659
Qy 3626 ATAGGGTTCCCGGCACATTTCCCGGAAAAAGTCCAC 3662
Db 3660 ATAGGGTTCCCGGCACATTTCCCGGAAAAAGTCCAC 3696

RESULT 4

US-09-503-222-6
; Sequence 6, Application US/09503222
; Patent No. 6265548
; GENERAL INFORMATION:
; APPLICANT: Pavlakis, George N.
; APPLICANT: Gaitanaris, George A.
; APPLICANT: Stauber, Roland H.
; APPLICANT: Vournakis, John N.
; TITLE OF INVENTION: Mutant Aequorea victoria fluorescent
; TITLE OF INVENTION: Proteins Having Increased Cellular Fluorescence
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/503,222
; FILING DATE:

```
;
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646, 538
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 015280-249000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3699 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: - 1..3699
; LOCATION: 1..3699
; OTHER INFORMATION: /note= "pBSGFP"
;
US-09-503-222-6

Query Match          71.6%; Score 2620.6; DB 3; Length 3699;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 3078; Conservative 0; Mismatches 579; Indels 40; Gaps 7;

QY 3 AAATTGTAAGCTTAAATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTCAT 62
Db 3 AAATTGTAAGCTTAAATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTCAT 62
QY 63 TTTTAAACCAATAGCGGAAATCGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGA 122
Db 63 TTTTAAACCAATAGCGGAAATCGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGA 122
QY 123 TAGGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACATATAAAGAACGTGGACTCCA 182
Db 123 TAGGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACATATAAAGAACGTGGACTCCA 182
QY 183 AGTCAAAGGGCGAAACACCGTCTATCAGGGCGATGGCCCACTAGTGAAACATCACCTT 242
Db 183 AGTCAAAGGGCGAAACACCGTCTATCAGGGCGATGGCCCACTAGTGAAACATCACCTT 242
QY 243 AATCAAGTTTTTGGGTCGAGTCCGCTTAAGAGCACTAAATCGGAACCTTAAGGGAGCC 302
Db 243 AATCAAGTTTTTGGGTCGAGTCCGCTTAAGAGCACTAAATCGGAACCTTAAGGGAGCC 302
QY 303 CCCGATTTAGAGCTTTGACGGGAAAGCCGCGAAACGTCGCGAAGAAAGGAAAGAAAG 362
Db 303 CCCGATTTAGAGCTTTGACGGGAAAGCCGCGAAACGTCGCGAAGAAAGGAAAGAAAG 362
QY 363 CGAAAGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACAGTGGCGGTAAACCAACA 422
Db 363 CGAAAGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACAGTGGCGGTAAACCAACA 422
QY 423 CACCCGCGCGCTTAATGCGCGCTACAGGGCGCTC---CCATTGCGCATTCAGGCTGC 479
Db 423 CACCCGCGCGCTTAATGCGCGCTACAGGGCGCTC---CCATTGCGCATTCAGGCTGC 482
QY 480 GCAACTGTGGGAAGGGGATCGGTGCGGGCTCTTCGCTATTATACCCAGCTGGCGAAAG 539
Db 483 GCAACTGTGGGAAGGGGATCGGTGCGGGCTCTTCGCTATTATACCCAGCTGGCGAAAG 542
QY 540 GGGGATGTCTCAAGGCGATTAAGTTGGGTAAACCGCAGGGTTTTCCAGTCACAGCTT 599
Db 543 GGGGATGTCTCAAGGCGATTAAGTTGGGTAAACCGCAGGGTTTTCCAGTCACAGCTT 602
QY 600 GTAAACGACGCCAGTGTAGCGCGCTCTGTTTCATTCAGCTTTTGAACCCGTGGAGGAC 659
Db 603 GTAAACGACGCCAGTGTAGTTGTAAT--ACGACTCACTATAGGGCGGAATTGGGTACCGG 660
QY 660 GGCAGACTCGCGGTGCAAAATGTGTTTACAGCGTGTAGGAGCAGATGAAGATGCTCGACA 719
```

```
Db 661 GCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATGATCCTTATTTGTATAGTTCAATC 720
QY 720 CGCTGCA---GAACAGCGAGCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTACG 776
Db 721 CATGCCATGTGTAAATCCAGCAGCTGTTTCAAACTCAAGAAAGGACCATGTGCTCTCTTT 780
QY 777 TTAAGATAAATCATGCGTAAATTCAGCATGGGATCTGTAATACGACTCATCATAGGCG 836
Db 781 TTCGTTGGGATCTTTCGAAAGGGCAGATTGTGTGACAGGTA-ATGGTTGTCTGGTAAAA 839
QY 837 GAATTGGGTACCGGCCCCCCCTCGAGGTGCAAGGTATCGATAAGCTTGAATCGAAATTC 896
Db 840 GGACAGGGCCATCGCCAATTGGAGTATTGTTGATAATGCTCTGCTAGTTGAACGCTTC 899
QY 897 CTGACAGCCGGGGGATCCACTAGTT---CTAGAGCGGCCGCCACCGCGGTGGAGCTCCA 952
Db 900 CATCTTCAATGTGTGCTAAATTTTGAAGTTAACTTTGATCCCACTTTTGTGTTGTCG 959
QY 953 GCTTTGTTCCTTTAGTGAGGGTTAAATTAGATCCCATCGCTCAATTTTACGACACTAT 1012
Db 960 CCATGATGTATACATTGTGTGATTAGTTTGTATTCCAATTTGTGTCCAAGATGTTTC 1019
QY 1013 CTTTCTAGGGTTAATCTAGCTGCAATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAG 1072
Db 1020 CATCTTCTTTAAATCAATACCTTTTAACTCGATTCTATTAAACAGGGGTATCACCTCAA 1079
QY 1073 TCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGAGGAAAGACCGCTGCTTTCCCG 1132
Db 1080 ACTTGACTTCAGCAGCTGTCTGTAGTTCCCGTCATCTTTGAAATAATAGTTCCTTCT 1139
QY 1133 CGAGGTTCAAGCGGCATCGAAAGAGTTTTCGCGAGGATGACTGCTGCTGATTCAGTTGA 1192
Db 1140 GTACATAACCTTCGGGCATGGCCTCTTTGAAAGAGTCAATCGCTTTCATATGATCOGG 1199
QY 1193 GCGAAACGCAAGTTTACCATGATGATTTCGGAAAGTGTGGCATGCGACGCTTTTCCGAC 1252
Db 1200 ATCTTGAAGAGCATTTGAACACCAATAAGAGAAAGTAGTGACAAGTGTGGCCATGGAAC 1259
QY 1253 TGAAGTGTGTTTCAGGCCACTGGGATACCAATTCGTCGGGCTTTTTCGGACACAGTT 1312
Db 1260 GTAGTTTTCAGTAGTGTGCAATAAATTTAAGGGTAAGTTTTCGGTATGTGTGATCACCT 1319
QY 1313 CCGGATGTGACGCCGGAAGCGCATCAGCAACCCGAAACAATACCGCGCAGACAGCCGGAAC 1372
Db 1320 CACCTCTCCACTGACAGAAATTTGTGCCCATTAACATCACCATCTAATTAACAAGAA 1379
QY 1373 CCGTGCC-----GGTGTGCAGATTAAATGACAGCGGTGCGCGCTGGGATAT 1418
Db 1380 TTGGACAACTCCAGTGAAGAGTTCTTCTCTTTGCTAGCCATTTCTTGGCGGATCGAAT 1439
QY 1419 TACGTCAGGAGGACGGGTATCTCGGCTGGATGCCGAGAAATGACATGGATACCCCGT 1478
Db 1440 TCCTGACGCCCGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGTCCACG 1499
QY 1479 GAGTTACCCCGCGG-----CGGCTTTGGCGTAAATCATGTGCTATAGCTGTTT 1525
Db 1500 TTTTGTCCCTTATAGTGGGGTTAATTCGAGCTTGGCGTAAATCATGTGCTATAGCTGTT 1559
QY 1526 CTTGTGAAATTTGTTATCCGCTCAAAATTCACAAACATACAGAGCCGGAAGCAATAAG 1585
Db 1560 CTTGTGAAAAATTTGTTATCCGCTCAAAATTCACAAACATACAGAGCCGGAAGCAATAAG 1619
QY 1586 TGTAAAGCCTGGGTGCTTAATGAGTGAAGTAACTCACATTAATTTGCTGGCTCACCTG 1645
Db 1620 TGTAAAGCCTGGGTGCTTAATGAGTGAAGTAACTCACATTAATTTGCTGGCTCACCTG 1679
QY 1646 CCGCTTTTCAGTGGGAAACCTGTGTCGCCAGCTGCAATTAATGAATGGCCAAACGCGCG 1705
Db 1680 CCGCTTTTCAGTGGGAAACCTGTGTCGCCAGCTGCAATTAATGAATGGCCAAACGCGCG 1739
QY 1706 GGGAGAGCGGTTTGTGCTGTTGGGGCTCTTCCGCTTCTCGCTCACCTGCTCGCTGGCG 1765
```

1740	GGAGAGGCGGTTTGGCGTATTTGGCGCTCTTTCGCTTCTTCGCTCACTGACTCGCTGCGC	1799
Qy	TCGGTTCGTTTCGCGTGGCGGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGCGTTATCC	1825
	TCGGTTCGTTTCGCGTGGCGGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGCGTTATCC	1859
	ACAGATACAGGGGATTAACGACAGGAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGG	1885
Db	ACAGAAATACAGGGATTAACGACAGGAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGG	1919
	AACCGTAAAAAGGCGCGTTTCTGGCGTTTTCATAGAGCTCCGCGCCCTGACGAGCAT	1945
	AACCGTAAAAAGGCGCGTTTCTGGCGTTTTCATAGAGCTCCGCGCCCTGACGAGCAT	1979
Qy	CACAAAATACGACGCTCAAGTCAAGAGTGCAGAACCCGACAGGACTATATAAGATACCCAG	2005
	CACAAAATACGACGCTCAAGTCAAGAGTGCAGAACCCGACAGGACTATATAAGATACCCAG	2039
	GCCTTTCCCGCTGGAAGCTCCCTCGTCGCTCTCTCTGTTCCGAGCCCTGCCGCTTACCGGA	2065
Db	GCCTTTCCCGCTGGAAGCTCCCTCGTCGCTCTCTCTGTTCCGAGCCCTGCCGCTTACCGGA	2099
	TACCTGTCGCGCTTTCCTCTCGGGAAGCGTGGCGCTTTCATAGCTCACGCTGTAGG	2125
	TACCTGTCGCGCTTTCCTCTCGGGAAGCGTGGCGCTTTCATAGCTCACGCTGTAGG	2159
Qy	TATCTCAGTTTCGGTGTAGTTCGTTTCGCTCCAAAGCTGGGCTGTGTGACGAACCCCGCTT	2185
	TATCTCAGTTTCGGTGTAGTTCGTTTCGCTCCAAAGCTGGGCTGTGTGACGAACCCCGCTT	2219
	CAGCCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCCGGTAAAGACAC	2245
Db	CAGCCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCCGGTAAAGACAC	2279
	GACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAAGCAGCGAGGTATGTAGGC	2305
	GACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAAGCAGCGAGGTATGTAGGC	2339
Qy	GGTGCTACAGAGTTCTTGAAGTGGTGCCCTAACTTACGGCTACACTAGAAGGACAGATTTT	2365
	GGTGCTACAGAGTTCTTGAAGTGGTGCCCTAACTTACGGCTACACTAGAAGGACAGATTTT	2399
	GGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCC	2425
Db	GGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCC	2459
	GGCAAAACAAACCCCGCTGGTAGCGGTGTTTTTGTGTTTGCAGCAGCAGATTAACGCGC	2485
	GGCAAAACAAACCCCGCTGGTAGCGGTGTTTTTGTGTTTGCAGCAGCAGATTAACGCGC	2519
Qy	AGAAAAAAGGATCTCAAGAAGATCCCTTGATCTTTTCTACGGGTCTGACGCTCAGTGG	2545
	AGAAAAAAGGATCTCAAGAAGATCCCTTGATCTTTTCTACGGGTCTGACGCTCAGTGG	2579
	AACGAAATCTCAGGTTAAGGATTTTGGTCAAGATTAATCAAAAAGGATCTTCACTAG	2605
Db	AACGAAATCTCAGGTTAAGGATTTTGGTCAAGATTAATCAAAAAGGATCTTCACTAG	2639
	ATCGTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATAGTAAACTTGG	2665
	ATCGTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATAGTAAACTTGG	2699
Qy	TCGACAGTTTACCAATCGCTTAACTCAGTGAAGCAGCTATCTCAGCGATCTCTCTATTTCGT	2725
	TCGACAGTTTACCAATCGCTTAACTCAGTGAAGCAGCTATCTCAGCGATCTCTCTATTTCGT	2759
	TCATCCAATAGTTGCTGACTCCCGCTCGGTAGATAACTACGATACGGGAGGCTTACCA	2785
Db	TCATCCAATAGTTGCTGACTCCCGCTCGGTAGATAACTACGATACGGGAGGCTTACCA	2819
	TCGTGGCCCCAGTGTGCAATGATACCGGAGACCCAGCGCTCACCGGCTCAGATTTATCA	2845
	TCGTGGCCCCAGTGTGCAATGATACCGGAGACCCAGCGCTCACCGGCTCAGATTTATCA	2879

RESULT 5

```

US-09-890-848-16
; Sequence 16, Application US/09890848
; Patent No. 6573053
; GENERAL INFORMATION:
; APPLICANT: Firth, Greg
; APPLICANT: Ogedra, Rajesh Muru
; TITLE OF INVENTION: Analysis Method
; FILE REFERENCE: P39903
; CURRENT APPLICATION NUMBER: US/09/890,848
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: PCT/G800/00346
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: EP 93300873.9
; PRIOR FILING DATE: 1999-02-05

```

; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 5277
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: plasmid
US-09-890-848-16

Query Match 67.4%; Score 2469.4; DB 4; Length 5277;
Best Local Similarity 82.3%; Pred. No. 0;
Matches 3014; Conservative 0; Mismatches 436; Indels 213; Gaps 7;

```
QY 3 AAATTGTAAGCGTTAAATTTTGTAAAAATTCGCGTTAAATTTTGTAAAAATCAGTCAAT 62
DB 1239 AAATTGTAACGTTAAATTTTGTAAAAATTCGCGTTAAATTTTGTAAAAATCAGTCAAT 1298
QY 63 TTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTTATAAATCAAAAGAAATAGACCGAGA 122
DB 1299 TTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTTATAAATCAAAAGAAATAGACCGAGA 1358
QY 123 TAGGTTGAGTGTGTTCCAGTTTGGAAACAAGTCCACTATTAAGAAACGTGGACTCCA 182
DB 1359 TAGGTTGAGTGTGTTCCAGTTTGGAAACAAGTCCACTATTAAGAAACGTGGACTCCA 1418
QY 183 AGCTCAAAAGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCT 242
DB 1419 AGCTCAAAAGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCT 1478
QY 243 AATCAAGTTTTTGGGTGTCAGGTGCGGTAAAGCACTAAATCGGAACCCCTAAAGGGAGCC 302
DB 1479 AATCAAGTTTTTGGGTGTCAGGTGCGGTAAAGCACTAAATCGGAACCCCTAAAGGGAGCC 1538
QY 303 CCCGATTTAGACTTTGACGGGGAAGCCGCGCAACGTCGCGGAAAGGGAAGAAAG 362
DB 1539 CCCGATTTAGACTTTGACGGGGAAGCCGCGCAACGTCGCGGAAAGGGAAGAAAG 1598
QY 363 CGAAAGGACGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGGTAAACCA 422
DB 1599 CGAAAGGACGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGGTAAACCA 1658
QY 423 CACCCGCGCGCTTAATGCGCGCTACAGGGCGCGTC---CCATTGCGCATTCAGGCTGC 479
DB 1659 CACCCGCGCGCTTAATGCGCGCTACAGGGCGCGTC---CCATTGCGCATTCAGGCTGC 1718
QY 480 GCAACTGTTGGGAAGGCGATCGGTGCGGGCTCTTCGCTATTACGCCAGCTGGCGAAG 539
DB 1719 GCAACTGTTGGGAAGGCGATCGGTGCGGGCTCTTCGCTATTACGCCAGCTGGCGAAG 1778
QY 540 GGGGATGTCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCAGTCCAGCAGCTT 599
DB 1779 GGGGATGTCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCAGTCCAGCAGCTT 1838
QY 600 GTAAACGACGCGCAGTGAAGTGGTCAATTCAGCTTTTGAACCCCGTGGAGGACG 659
DB 1839 GTAAACGACGCGCAGTGAAGTGGTCAATTCAGCTTTTGAACCCCGTGGAGGACG 1898
QY 660 GGCAGACTCGGGTCAAAATGTTTACAGCGTGTAGGAGAGATGAAGTGTCTGACA 719
DB 1899 AATAAGATTCTACAATACTAGCTTTTATGGTTATGAAGAGGAAAAATTTGGCAGTAACT 1958
QY 720 CGCTCAGAACACGCGCTAGATTAACCTAGAAAGATAATCATATTGTGAGTACGTTA 779
DB 1959 GGGCCCAACAAC-----TTCAATGAACGAATCAAAATTAACCAACATAGGATGATA 2010
QY 780 AAGATAATCATCGTAAATTAAGCATGGGATCTGTAATACGACTCACTATAGGGCGAA 839
DB 2011 ATGCGATTAGTTTTTATGCTTATTTCTGGGGTAATTAATCAGCGAAGCGAT---GATT 2066
QY 840 TTGGGTACGGGCCCCCTCGAGGTCGACGGTATCGATAGCTTGATATCGAATTCCTG 899
```

```
DB 2067 TTTGATCTATTAAACAGATATATAAATGCAAAAAATGCAATTAACCACTTTAACTAATCTTT 2126
QY 900 CAGCCCGGGGATCCACTAGTTCTAGAGCGGCGCCACCGCGGTGAGCTCCAGACTTTTG 959
DB 2127 CAACATTTTCGGT-----TTGTATTAC 2148
QY 960 TTCCCTTTAGTGAGGGTTAAATTTAGATCCCATGCGTCAATTTTACGAGACTATCTTTCTA 1019
DB 2149 TTCTTATTCAATGTAAATAAAGATATCAACATCTAGAAATTCATGGAATCGATCGGCGG 2208
QY 1020 GGGTTAATCTAGTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCG 1079
DB 2209 CAATTCATCGATTAGACTAGTCTAGATATCATGTAATAGTTATGTCACGCTTACATCCA 2268
QY 1080 CCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGACCGCGTCTTTTCCCGCGAGGTT 1139
DB 2269 CGCC----- 2272
QY 1140 GAAGCGCATGGAAGAGTTTTCGAGGATGATCTCTGTCATTTGAGCTTGAAGCGTGAAGAAA 1199
DB 2273 -----CTCCCCCACAATCCGCTCTTAACCGAAAA 2300
QY 1200 CGCAGCTTACCATGATGATTCGGGAAGGTGGGCCATGCAAGCCTTTAAACGTTGNACTG 1259
DB 2301 GGAAGGATGAGACAACCTGAGTCTAGTCCCTATTTTATTTTATAGTTATGTAGT 2360
QY 1260 TTGCTTACGGCCACCTCGGATACCAATTCGTCGGGTCTTTCCGGACACAGTTCCGGATG 1319
DB 2361 ATTAAGAAGCTTATTATTTTCAAAATTTTCTTTT----- 2400
QY 1320 CTCAGCCCGAAGCGCATCAGCAACCCGAAACATACCGGCGACAGCCGGAATTCGCGTGC 1379
DB 2401 -----CTGTACAGACGCGTGTACGCATGTAAACATATATACT 2435
QY 1380 GGTGTGAGATTAAATGACAGCGGTGCGGCGTGGGATATTAGCTCAGCGAGACCGGTAT 1439
DB 2436 GAAACCTTGTGTGAGAAGGTTTGGGACGCTCGAA----- 2471
QY 1440 CTGCTCGATGCCGAGAAATGGACATACCCCGTGTAGTTACCCGCGGGCGCGCT 1499
DB 2472 ---GGCTTTAATTTGCGGTACCCAGCTTTTGTTCCTTTAGTGAGGTTAATTCGAGCT 2528
QY 1500 TGGCGTAATCATGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCAAAATCCAC 1559
DB 2529 TGGCGTAATCATGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCAAAATCCAC 2588
QY 1560 ACAACATACGAGCGGAGAGCATAAAGTCTAAAGCTGGGTGCTTAATGAGTGAGCTAAC 1619
DB 2589 ACAACATAGGAGCGGAGAGCATAAAGTCTAAAGCTGGGTGCTTAATGAGTGAGCTAAC 2648
QY 1620 TCACATTAATTTGCGTTGCGCTCACTGCGCGCTTTCCAGTCGGGAAACCTGTCGTCGCAGC 1679
DB 2649 TCACATTAATTTGCGTTGCGCTCACTGCGCGCTTTCCAGTCGGGAAACCTGTCGTCGCAGC 2708
QY 1680 TGCAATTAATGAAATCGGCAACCGCGGGAGAGCGGTTTGGCTATTGGGCGCTCTTCG 1739
DB 2709 TGCAATTAATGAAATCGGCAACCGCGGGAGAGCGGTTTGGCTATTGGGCGCTCTTCG 2768
QY 1740 CTTCTCGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1799
DB 2769 CTTCTCGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2828
QY 1800 ACTCAAGCGGTAAATACGTTATCCACAGAAATCAGGGGATAACCGCAGGAAAGACATGT 1859
DB 2829 ACTCAAGCGGTAAATACGTTATCCACAGAAATCAGGGGATAACCGCAGGAAAGACATGT 2888
QY 1860 GAGCAAAAGGCGCAAAAGGCGGAAACCGTTAAAGGCGCGCTGCTGCTGCGTTTTC 1919
DB 2889 GAGCAAAAGGCGCAAAAGGCGGAAACCGTTAAAGGCGCGCTGCTGCTGCGTTTTC 2948
QY 1920 ATAGCTCCGCGCCCTCGAGGATCAAAATTCGACGCTCAAGTCAGAGTGGCGAA 1979
DB 2949 ATAGCTCCGCGCCCTCGAGGATCAAAATTCGACGCTCAAGTCAGAGTGGCGAA 3008
```

1980 ACCGACAGGACTATAAGATACACAGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTC 2039
1981 |||||
3009 ACCGACAGGACTATAAGATACACAGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTC 3068
3010 |||||
2040 CTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGG 2099
2041 |||||
3069 CTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGG 3128
3070 |||||
2100 CGCTTTCTCANTAGCTCAGCGCTAGGTATCTCAGTTCCGTTGAGTGGTGGTTCGCTCAAGC 2159
2101 |||||
3129 CGCTTTCTCAATGCTCAGCGCTAGGTATCTCAGTTCCGTTGAGTGGTGGTTCGCTCAAGC 3188
3130 |||||
2160 TGGGCTGTGTGACAGAACCGCCCGTTTCCGCGGACCGCTGCGCTTATCGGTTAACTATC 2219
2161 |||||
3189 TGGGCTGTGTGACAGAACCGCCCGTTTCCGCGGACCGCTGCGCTTATCGGTTAACTATC 3248
3190 |||||
2220 GTCTTCAGTTCACACCGGTAAGACACGACTTATCGCACCTGGCAGCAGCCACTGGTAACA 2279
2221 |||||
3249 GTCTTCAGTTCACACCGGTAAGACACGACTTATCGCACCTGGCAGCAGCCACTGGTAACA 3308
3250 |||||
2280 GGATTAGCAGGAGGAGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAAC 2339
2281 |||||
3309 GGATTAGCAGGAGGAGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAAC 3368
3310 |||||
2340 ACGGCTACACTAGACGACGATTTGGTATCTGCGCTCTGCTGGAAGCAGTTACCTTCG 2399
2341 |||||
3369 ACGGCTACACTAGACGACGATTTGGTATCTGCGCTCTGCTGGAAGCAGTTACCTTCG 3428
3370 |||||
2400 GAAAAAGAGTTGTAGCTCTTGATCGCGCAACCAACACCGCTGCTAGCGGTGGTTTTT 2459
2401 |||||
3429 GAAAAAGAGTTGTAGCTCTTGATCGCGCAACCAACACCGCTGCTAGCGGTGGTTTTT 3488
3430 |||||
2460 TTGTTTTCAGCAGCAGATTTACCGCGAGAAAAAGGATCTCAAGAGATCTCTTTCATCT 2519
2461 |||||
3489 TTGTTTTCAGCAGCAGATTTACCGCGAGAAAAAGGATCTCAAGAGATCTCTTTCATCT 3548
3490 |||||
2520 TTTCTACGGGCTCAGCTCAGTGGAACGAAACTCAGCTTAAAGGATTTGGTCAAGA 2579
2521 |||||
3549 TTTCTACGGGCTCAGCTCAGTGGAACGAAACTCAGCTTAAAGGATTTGGTCAAGA 3608
3550 |||||
2580 GATTATCAAAAAGGATCTTCACTAGATCTCTTAAATTAATAAATCAAGTTTAAATCAA 2639
2581 |||||
3609 GATTATCAAAAAGGATCTTCACTAGATCTCTTAAATTAATAAATCAAGTTTAAATCAA 3668
3610 |||||
2640 TCTAAAGTATATAGTAAACTTGTCTCAGAGTTACCAATGCTTAAATCAGTAGGACAC 2699
2641 |||||
3669 TCTAAAGTATATAGTAAACTTGTCTCAGAGTTACCAATGCTTAAATCAGTAGGACAC 3728
3670 |||||
2700 CTATCTCAGGATCTCTCTATTTTCTATTCATTCATTCAGTTCGCTGCTGCTGCTGCTG 2759
2701 |||||
3729 CTATCTCAGGATCTCTCTATTTTCTATTCATTCATTCAGTTCGCTGCTGCTGCTGCTG 3788
3730 |||||
2760 TAACTACGATACGGAGGCTTACCATCTGGCCCGAGTCTGCAATGATACCGGAGCC 2819
2761 |||||
3789 TAACTACGATACGGAGGCTTACCATCTGGCCCGAGTCTGCAATGATACCGGAGCC 3848
3790 |||||
2820 CACGCTCACGGCTCCAGATTTATCAGCAATAAACCCAGCAGCGGAGGCGGAGCGCA 2879
2821 |||||
3849 CACGCTCACGGCTCCAGATTTATCAGCAATAAACCCAGCAGCGGAGGCGGAGCGCA 3908
3850 |||||
2880 GAAGTGGTCTGCAACTTTATCGGCTCCATCCAGTCTTAAATTTGTCGGGAAGCTTA 2939
2881 |||||
3909 GAAGTGGTCTGCAACTTTATCGGCTCCATCCAGTCTTAAATTTGTCGGGAAGCTTA 3968
3910 |||||
2940 GAGTAAAGTATTCGCGCAGTAAAGTTTGGCAACCGTTGCTGCAATGCTACAGGATCG 2999
2941 |||||
3969 GAGTAAAGTATTCGCGCAGTAAAGTTTGGCAACCGTTGCTGCAATGCTACAGGATCG 4028
3970 |||||
3000 TGGTGTACGCTCGTCTGTTGGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGC 3059
3001 |||||
4029 TGGTGTACGCTCGTCTGTTGGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGC 4088
4030 |||||

3060 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCG 3119
3061 |||||
4089 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCG 4148
4090 |||||
3120 TTGTGAGAAAGTAAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATT 3179
3121 |||||
4149 TTGTGAGAAAGTAAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATT 4208
4150 |||||
3180 CTCCTTACTCTCATGCAATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTACTCAACAAGT 3239
3181 |||||
4209 CTCCTTACTCTCATGCAATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTACTCAACAAGT 4268
4210 |||||
3240 CATTCTGAGAAAGTGTATGCGCGCAGTGTCTTTTCCCGCGCTCAATACGGGATA 3299
3241 |||||
4269 CATTCTGAGAAAGTGTATGCGCGCAGTGTCTTTTCCCGCGCTCAATACGGGATA 4328
4270 |||||
3300 ATACCGCGCACATAGCAGAACTTTAAAAAGTGTCTCATCTGGAAAAAGTCTTCTTCGGGCG 3359
3301 |||||
4329 ATACCGCGCACATAGCAGAACTTTAAAAAGTGTCTCATCTGGAAAAAGTCTTCTTCGGGCG 4388
4330 |||||
3360 GAAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCCGATGTAACCCACTCGTGCAC 3419
3361 |||||
4389 GAAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCCGATGTAACCCACTCGTGCAC 4448
4390 |||||
3420 CCAACTGATCTTCAAGGATCTTACCGCTGTGAGATCCAGTTCCGATGTAACCCACTCGTGCAC 3479
3421 |||||
4449 CCAACTGATCTTCAAGGATCTTACCGCTGTGAGATCCAGTTCCGATGTAACCCACTCGTGCAC 4508
4450 |||||
3480 GCAAAATCCGCCAAAAAGGAAATAAGGGCGACACGGAATGTTGAAATCTCATACTCT 3539
3481 |||||
4509 GCAAAATCCGCCAAAAAGGAAATAAGGGCGACACGGAATGTTGAAATCTCATACTCT 4568
4510 |||||
3540 TCCTTTTCAATTAATTGAAGCAATTAATCAGGTTTATGTTCTCATGAGCGGATACATAT 3599
3541 |||||
4569 TCCTTTTCAATTAATTGAAGCAATTAATCAGGTTTATGTTCTCATGAGCGGATACATAT 4628
4570 |||||
3600 TTGAATGATTTAGAAAAATAAACAAATAGGGTTCCCGCGCACTTCCCGGAAAGTGC 3659
3601 |||||
4629 TTGAATGATTTAGAAAAATAAACAAATAGGGTTCCCGCGCACTTCCCGGAAAGTGC 4688
4630 |||||
3660 CAC 3662
3661 |||||
4689 CAC 4691
4690 |||||

RESULT 6
US-09-402-266B-21
; Sequence 21, Application US/09402266B
; Patent No. 6537767
; GENERAL INFORMATION:
; APPLICANT: HINNEN, Albert
; APPLICANT: HEGEMANN, Johannes
; APPLICANT: MUNDER, Thomas
; APPLICANT: SCHUSTER, Tilmer
; APPLICANT: FELDMMANN, Horst
; APPLICANT: KRAMER, Wilfried
; APPLICANT: ZIMMERMANN, Friedrich
; APPLICANT: ENTIAN, Karl-Dieter
; TITLE OF INVENTION: Process for Screening Antimycotically Active Substances
; FILE REFERENCE: 38005-0094
; CURRENT APPLICATION NUMBER: US/09/402,266B
; CURRENT FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: PCT/EP98/01904
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: DE 19713572.2
; PRIOR FILING DATE: 1997-04-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 3956
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:


```
; NAME/key: misc_feature
; LOCATION: (0..7)
; OTHER INFORMATION: Synthetic sequence
US-09-402-266B-21

Query Match      66.9%; Score 2448.8; DB 4; Length 3956;
Best Local Similarity 79.4%; Pred. No. 0;
Matches 3136; Conservative 0; Mismatches 522; Indels 294; Gaps 8;

QY 5 ATTGTAAGCGTTAAATATTTTGTAAATTCGGGTAAATTTTGTAAATCAGCTCATTT 64
D  |||||
D  1 ATTGTAAGCGTTAAATATTTTGTAAATTCGGGTAAATTTTGTAAATCAGCTCATTT 60
QY 65 TTTAAACCAATAGCCGGAATCCGCAAAATCCCTTATAATCAAAAGATAGACCGAGATA 124
D  |||||
D  61 TTTAACCAATAGCCGGAATCCGCAAAATCCCTTATAATCAAAAGATAGACCGAGATA 120
QY 125 GGGTTGAGTGTGTTCCAGTTTGGAAACAAGATCCACTATTAAGAAAGCGTGAATCCCAAC 184
D  |||||
D  121 GGGTTGAGTGTGTTCCAGTTTGGAAACAAGATCCACTATTAAGAAAGCGTGAATCCCAAC 180
QY 185 GTCAAGGGCGGAACCGCTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCTAA 244
D  |||||
D  181 GTCAAGGGCGGAACCGCTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCTAA 240
QY 245 TCAAGTTTTCGGGTTCGAGTGCCTTAAGCACTAAATCGGAACCTTAAAGGAGAGCCCC 304
D  |||||
D  241 TCNAGTTTTCGGGTTCGAGTGCCTTAAGCACTAAATCGGAACCTTAAAGGAGAGCCCC 300
QY 305 CGATTTAGAGCTTGACGGGGAAGCGCGGAACGTTGGGAGAGAAAGGGAAGAAAGCG 364
D  |||||
D  301 CGATTTAGAGCTTGACGGGGAAGCGCGGAACGTTGGGAGAGAAAGGGAAGAAAGCG 360
QY 365 AAAGAGCGGGCGCTAGGGCGCTGCGCAAGTGTAGCGGTACAGCTGGGCTAACCCACCA 424
D  |||||
D  361 AAAGAGCGGGCGCTAGGGCGCTGCGCAAGTGTAGCGGTACAGCTGGGCTAACCCACCA 420
QY 425 CCCGCGCGCTTAATGCGCGCTACAGGGCGGCTCCCATTCGCCATTACGGCTGCGCAAC 484
D  |||||
D  421 CCCGCGCGCTTAATGCGCGCTACAGGGCGGCTCCCATTCGCCATTACGGCTGCGCAAC 480
QY 485 TGTGGGAAGGCGATCGGTGCGGCGCTCTTCGCTATTACGCCAGCTGCGCAAGGGGGA 544
D  |||||
D  481 TGTGGGAAGGCGATCGGTGCGGCGCTCTTCGCTATTACGCCAGCTGCGCAAGGGGGA 540
QY 545 TGTGCTGCAAGCGGATTAAGTTGGGTAAAGCCAGGGTTTCCAGTACGAGCTGTATAA 604
D  |||||
D  541 TGTGCTGCAAGCGGATTAAGTTGGGTAAAGCCAGGGTTTCCAGTACGAGCTGTATAA 600
QY 605 ACGAGCGGCAGTGAGCGGCGCTCGTTTCATTACGTTTGTGAACCGGTGGAGGACGGGAG 664
D  |||||
D  601 ACGAGCGGCAGTGAGCGGCGCTTAATACGACTCACTATAGGGCGAATTGGAGCTCCACGC 660
QY 665 ACTCGCGGTGCAAAATGTGTTTACAGCGGTATGGAGC----- 701
D  |||||
D  661 GGTGGCGCTCTAGAACTAGTGATCGGCAATGGAGGCCACACCGCATAGATCGGCAAGT 720
QY 702 ----- 701
D  721 GCACAAACAATACTTAAATAAATACTACTAGTAATAACCTTATTTCTTAGCAATTTTGAC 780
QY 702 -----AGATGAAGATGCTCGACACGCTGAGAACACGACGAGCTAGATT 743
D  |||||
D  781 GAAATTTGCTATTTTGTAGAGCTTTTACACCATTTGTCTCCACCTCCGCTTACATC 840
QY 744 AACCCTAGAAAGATATCATATTTGACGTAAGTTAAAGATATCATGCGT----- 794
D  |||||
D  841 AACACCAATAACGCCATTTAACTAAGCGCATCAACCAATTTCTGGCGTCAGTCCACC 900
QY 795 -----AAAATTGACGATGGATCTGTAATACGACTCACTATATAGGGGGAATTT--- 841
D  |||||
D  901 AGCTAACATAAAATGTAAGCTTTTCGGGCGCTCTCTGCTTCCAAACCCAGTCAGAAATCGA 960
```

```
842 ----- 841
961 GTTCCAATCAAAAGTTCACTGTGCCACCTGTCTTGAATCAAAACAAAGGAATAAAACGA 1020
842 -----GGGTACCGGGCCCCCTCGAGGTCGACGGTATCGATAAGCTTGTATATCGAATTC 897
1021 ATGAGGTTTCTGTGAAGCTGCACGTAGTAGTATGTTGCAGTCTTTTGGAAATACGAGTCT 1080
898 TGCAGCCCCGGGGATCCACTAGTTCTTAGAGGGCGCCACCGCGTGGAGCTCCA----- 952
1081 TTTAATAACTGCAAAACCGAGAACTCTTGTATTCTTGGCCACGACTCATCTCCATCGAG 1140
953 --GCTTTTGTTCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGAGACT 1010
1141 TTGGACGATATCAATGCCGTAACTATGACAGAGCCAAACATCTCTCTTAGTGTGATT 1200
1011 ATCTTTC----- 1017
1201 ACGAAACACGCCAACCAAGTATTTCCGGAGTGCCTGAACCTATTTTATATGCTTTTACAG 1260
1018 ---TAGGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTC 1074
1261 ACTTGAATTTTCTTGAATTAACCGGGTCAATGTTCTCTTCTTATTGGGCACACATAT 1320
1075 ATCGCCCAAGCTGGCGCTATCTG-----GGCATCGGGAGGAAGAA 1115
1321 AATACCCAGCAAGTCAGCATCGGAATCTAGAGCACATTTCTGCGGCTCTGTCTCTGCAA 1380
1116 GCCGTGCTTTTCCGGGAGGTTGAAGGGCATGGAAGAGTTTTCGCGAGATGACTGC 1175
1381 GCCGAAACCTTTACCAATGGACAGAACTACCTGTGAATTAATAACAGACATACTCCA 1440
1176 TGCTCATTTGAGCTGTGAG-----CGAAAAACGACGTTTACCATGATGATTCGGG----- 1224
1441 AGCTCCCTTGTGCTTAATCAGCTATACTCAGGTGCTCAATAGTCACCAATGCCCTCC 1500
1225 -----AAGGTGTGGCCATCAGCGCTTTAAACGGTGAACCTTTCGTTTCAGGC 1270
1501 CTCCTTGGCCCTCTCCTCTTTTTCGACCGAAATTAATTTCTTAATCGCAAAAAAGAA 1560
1271 CACTGGGATACAGTTCTGTCGGCTTTTCGGGACACAGTTCCGGATGGTCAGCCGNA 1330
1561 AAGCTCCGGATCAAGATTGTACGTAAGGTGACAAGCTATTTTCAATAAAGAAATATCTTC 1620
1331 GCGCATCAGCAACCGGAACAATACCGCGACAGCGGAACTGCGCTGCGGTGTGCAGAT 1390
1621 CACTACTGCCATCTGCGCTATACTGCAAGTAGTACATATATTACGATGCTGTTCTATT 1680
1391 TAAATGACAGCGTGGCGCTGGGATATTACGTACAGGAGGACGGGTATCTCTGGCTGGAT 1450
1681 AAATGCGGCGCTGCAGGAATTCGATATCAAGCTTATCGATACCGCTCGACCTCGAGGGG 1740
1451 GCGCAGAAATGGACATGGATACCCCGTAGTTACCGCGGGCGGCTTGGCGTAAATCA 1510
1741 GCGCGGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTAAATTTGCGCGCTTGGCGTAAATCA 1800
1511 TGGTCATAGCTGTTTCTGTTGAAATTTGTTATCCGCTCACAATTCACACAAATACGA 1570
1801 TGGTCATAGCTGTTTCTGTTGAAATTTGTTATCCGCTCACAATTCACACAAATACGA 1860
1571 GCCGAAGCATAAAGTGTAAAGCTTGGGTGCTTAATGAGTGAGTAACTCACTCAATTAAT 1630
1861 GCCGAAGCATAAAGTGTAAAGCTTGGGTGCTTAATGAGTGAGTAACTCACTCAATTAAT 1920
1631 GGGTTGCGCTCACTGCCGCTTTCCAGTCGGGAACCTGTCGTGCGAGCTGCAATTAATGA 1690
1921 GGGTTGCGCTCACTGCCGCTTTCCAGTCGGGAACCTGTCGTGCGAGCTGCAATTAATGA 1980
1691 ATCGGCCAAACGCGGGGAGAGCGGTTTGGCTATTTGGCGCTCTTTCGCTTCCTCGCTC 1750
1981 ATCGGCCAAACGCGGGGAGAGCGGTTTGGCTATTTGGCGCTCTTTCGCTTCCTCGCTC 2040
1751 ACTGACTCGCTCGCTCGGTTCGTTTCGCTGCGGCGAGCGGTTATCAGCTCACTCAAAAGGCG 1810
```


;
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: DE 19713572.2
; PRIOR FILING DATE: 1997-04-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 4102
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: ()..()
; OTHER INFORMATION: Synthetic sequence
US-09-402-266B-20

Query Match 62.9%; Score 2305.2; DB 4; Length 4102;
Best Local Similarity 76.9%; Pred. No. 0;
Matches 3150; Conservative 0; Mismatches 508; Indels 440; Gaps 10;

QY 5 ATTGTAAGCGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCATTT 64
DB 1 ATTGTAAGCGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCATTT 60

QY 65 TTTAAACCAATAGCCGGAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACGAGATA 124
DB 61 TTTAAACCAATAGCCGGAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACGAGATA 120

QY 125 GGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTTAAAGAAACGTGACTCCAAAC 184
DB 121 GGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTTAAAGAAACGTGACTCCAAAC 180

QY 185 GTCAAGGCGGAAACCGTCTATCAGGCGGATGGCCCACTACGTGAACCATCACCCCTAA 244
DB 181 GTCAAGGCGGAAACCGTCTATCAGGCGGATGGCCCACTACGTGAACCATCACCCCTAA 240

QY 245 TCAAGTTTTTTGGGGTCGAGGTGCGGTAAAGCACTAAATTCGGAACCTTAAAGGAGCCCC 304
DB 241 TCAAGTTTTTTGGGGTCGAGGTGCGGTAAAGCACTAAATTCGGAACCTTAAAGGAGCCCC 300

QY 305 CGATTAGAGCTTGACGGGAAAGCGGCGACGTGGCGAGAAAGGGAAGGAAGG 364
DB 301 CGATTAGAGCTTGACGGGAAAGCGGCGACGTGGCGAGAAAGGGAAGGAAGG 360

QY 365 AAAGGAGCGGCGCTAGGCGCTGCGCAAGTGTAGCGGTCACTGCGGTGAACCAACCA 424
DB 361 AAAGGAGCGGCGCTAGGCGCTGCGCAAGTGTAGCGGTCACTGCGGTGAACCAACCA 420

QY 425 CCCGCCGCTTAAATGCGCGCTACAGGCGCGTCCCAATTCGCCAATTCAGGCTGCGCAAC 484
DB 421 CCCGCCGCTTAAATGCGCGCTACAGGCGCGTCCCAATTCGCCAATTCAGGCTGCGCAAC 480

QY 485 TGTGGGAAGGCGGATCGGTGCGGCGCTTTCGCTATTACGCCAGCTGGCGAAAGGGGA 544
DB 481 TGTGGGAAGGCGGATCGGTGCGGCGCTTTCGCTATTACGCCAGCTGGCGAAAGGGGA 540

QY 545 TGTGCTGCAAGCGGATTAAATTTGGGTAAACGACAGGTTTTCCAGTCAACAGTTGTAAA 604
DB 541 TGTGCTGCAAGCGGATTAAATTTGGGTAAACGACAGGTTTTCCAGTCAACAGTTGTAAA 600

QY 605 ACAGCGGCGAGTGAGCGGCGCTCGTTTCATTCAGTTTGTAAACCGGTGGAGGACGGGCG 664
DB 601 ACAGCGGCGAGTGAGCGGCGTAAATAGCACTCACTATAGGGCGAATTTGAGTCCACCGC 660

QY 665 ACTCGGGTGCAATGTGTTTTACAGCGTGATGGAGCAGATGAAGA----- 710
DB 661 GGTGCGCTCTAGAACTAGTGATCGGCGCATGGAGCGCTTCAAGAAATTAGCTTTTCAAT 720

QY 711 ----- 710
DB 721 TCAATTCATATTTTTTTTATTTCTTTTTTGTGATTTCCGTTCTTTGAAATTTTTT 780

QY 711 ----- 710

DB 781 ATTTCGGTAATCTCCGAACAGAAAGGAACGAAGGAGGACACAGACTTAGATTGGTAT 840
QY 711 ----- 710
DB 841 ATATACGCATATGTAGTGTGAAGAAACATGAATTTGCCAGTATTTCTTAACCCAACTGC 900
QY 711 --TGCTCGACACGCTGCAGAAACACGACGCTAGATTAAACCTAG----- 751
DB 901 ACAGAACAAAACATGACGAGAAACGAAGATAAATCATCTCGAAGCTACATATAAGGAAC 960
QY 752 -----AAAGA 756
DB 961 GTGCTGCTACTCATCTAGTCTCTGTTGCTGCTCAAGCTATTAAATATCATGCACGAAAAGC 1020
QY 757 TAATCATATTGTA-----CGTAGCTTAAAGATATCATGCGTAAATTTG 801
DB 1021 AAACAAACTGTGTGCTCTTCAATTGGATGTTGCTACCAACCAAGAAATTTCTTGGAGTTAGTTG 1080
QY 802 AGCATGGGATCTGTAATACGACTCACTATAGGGCGAATTTGGGTACCGGGCCCCCCTCG 861
DB 1081 AAGCATTAGTCCCAAAATTTGTTTACTTAAACACATGCTGGATATCTTGACTGATTTT 1140
QY 862 AGGTCGACGGTATCGATAAGC----- 882
DB 1141 CCATGGAGGCGACAGTTTAAAGCGCTTAAAGGCATTATCGCCAAGTACAATTTTTTACTCT 1200
QY 883 -----TTGATATCGAATTTCTCGAGCCCGGGGAT 912
DB 1201 TCGAAGACAGAAATTTCTGACATTTGTTAATACAGTCAAATTTGAGTACTCTCGGGGTG 1260
QY 913 CCACCTAGTTTCTAGAGCGCGCG-----CACCGCGGTGAGCTCCAGCTT 956
DB 1261 TATACAGATAGCAGATGGGCGACACATTACGAATGCACACGTTGTGTGGGCCCGGTA 1320
QY 957 TTGTTCCCTTTAGTGAGGG-----TTAA 979
DB 1321 TTGTTAGCGGTTTGAAGCAGCGCGGAGAAAGTAACAAGGAACCTAGAGCCCTTTTGA 1380
QY 980 TTAGATCCCATCGCTCAATTTTACGACACTATCTTTCTAGGGTTAACTAGCTCATCA 1039
DB 1381 TGTTAGCAGATTTGTTCGAAAGGCTCCCTATCTACTGGAGAAATATTAAGGCTACTG 1440
QY 1040 GGATCATATCTGTCGGG-----TCTTTTTTCGGCTCAGTCACTCCCAAGCTGCG 1089
DB 1441 TTGACATTTGCGAAGAGCGCAAGATTTTGTATCGGCTTTATTTGCTCAAAGAGACATGG 1500
QY 1090 GCTATCTGGGATCGGGAGGAAAGCCCGTGCCTTTTCCGCGAGGTTGAAGCGGAT 1149
DB 1501 GTGGAAGAGATGAAGGTTTACGATTTGTTGATTATGACACCCGCTGTGGTTTAGATGACA 1560
QY 1150 GGAAAGAGTTTGCAGGAGTCACTGCTGCTGATTTGAGGTTGAGGAAAAACGACGTTTA 1209
DB 1561 AGGGAGACGATTTGGGTCAAAGTATAGAACCGTGGATGATGTGTCTCTCAGGATCTG 1620
QY 1210 CCATGATGATTCGGGAAGGTTGTGGCCATGCACGCTTTA-----ACGGTG 1254
DB 1621 ACATTTATTTTGTGGAAGAGCACTATTTGCAAGGAAGGAGTCTAAGGTAGAGGGTG 1680
QY 1255 AACTGTTCTGTCAGGCCACCTGGGATACAGTT-----CGTCGCGGCTTTTCCGAGACACA 1309
DB 1681 AACGTTACAGAAAAAGCGCTGGGAAGCATATTTGAGAAGATGTCGGCCGACGAAAACTAAA 1740
QY 1310 GTTCCGGATGTCAGCCCGAGCGCATCAGCAACCCGCAACATACCGGCGAC-----AGC 1364
DB 1741 AAACGTGATTTAAGTAAATGATGATATCTAAACTCAAAATTAGAGCTTCAATTTAAT 1800
QY 1365 CGGAACCTCCGTCGCGGTGTCAGATTAATGACAGCGGTGGCGGCTGGGATATTACGTC 1424
DB 1801 TATATCAGTTATACCAATTTCTCATGTTTGGCGCGCTGCAGGAATTCGATATCAAGCT 1860
QY 1425 ACCGAGGACGGGTATCTCTGGCTGAGTGCAGGAGAAATGGAATGGAATACCGCTGAGTTA 1484
DB 1861 TATCGATACCGTTCGACCTTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCTTTAGTGAG 1920

1485 CCGGCGGGCGCGCTTGGCGTAATCATGATCATAGCTGTTTCTGTGTGAAATTTGTTATC 1544
1921 GGTAAATTGCGCGCTTGGCGTAATCATGATCATAGCTGTTTCTGTGTGAAATTTGTTATC 1980
1545 CGCTCACAATTCACACACATACGAGCGCGGAGCATAAAGTGTAAAGCTGGGTGCGCT 1604
1981 CGCTCACAATTCACACACATACGAGCGCGGAGCATAAAGTGTAAAGCTGGGTGCGCT 2040
1605 AATGAGTGAGCTAACTCACATTAATTTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAA 1664
2041 AATGAGTGAGCTAACTCACATTAATTTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAA 2100
1665 ACCTGTGCTGCGAGCTGCATTAATGAATGCGCCAAACGCGCGGGGAGAGCGGTTTCGGTA 1724
2101 ACCTGTGCTGCGAGCTGCATTAATGAATGCGCCAAACGCGCGGGGAGAGCGGTTTCGGTA 2160
1725 TTGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGCTGCGCTCGCGC 1784
2161 TTGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGCTGCGCTCGCGC 2220
1785 GAGCGGTATCAGCTCACTCAAGCGCGGTAAATACGTTTATCCACAGAAATCAGGGGATAACG 1844
2221 GAGCGGTATCAGCTCACTCAAGCGCGGTAAATACGTTTATCCACAGAAATCAGGGGATAACG 2280
1845 CAGGAAGAAATGTCAGCAAAAGGCGCAAAAGGCGCAAAAGGCGCAAAAGGCGCGCT 1904
2281 CAGGAAGAAATGTCAGCAAAAGGCGCGCAAAAGGCGCAAAAGGCGCGCT 2340
1905 TGCCTGCGCTTTCATAGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCA 1964
2341 TGCCTGCGCTTTCATAGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCA 2400
1965 GTCAGAGGTGGGAAACCCGACAGGACTATAAGATACACAGGCGTTTCCCGCTGGAAGCT 2024
2401 GTCAGAGGTGGGAAACCCGACAGGACTATAAGATACACAGGCGTTTCCCGCTGGAAGCT 2460
2025 CCTCTGCGCTCTCTGTTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTCTCC 2084
2461 CCTCTGCGCTCTCTGTTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTCTCC 2520
2085 CTTCCGGAGCGTGGCGCTTCTCATAGCTCAGCTGAGGTATCTCAGTTCGGTGTAG 2144
2521 CTTCCGGAGCGTGGCGCTTCTCATAGCTCAGCTGAGGTATCTCAGTTCGGTGTAG 2580
2145 TCGTTCGCTCCAGCTGGGCTGTGTCAGCAACCCCGCTTACGCCCGACCGCTGCGCT 2204
2581 TCGTTCGCTCCAGCTGGGCTGTGTCAGCAACCCCGCTTACGCCCGACCGCTGCGCT 2640
2205 TATCCGGTAACTATCGTCTTGAAGTCCAAACCCCGTAAAGACAGCTTTATCGCCATGGCAG 2264
2641 TATCCGGTAACTATCGTCTTGAAGTCCAAACCCCGTAAAGACAGCTTTATCGCCATGGCAG 2700
2265 CAGCCACTGGTAACAGATATGACAGAGCGAGGTATGTAGCGGTGTCTACAGTTCCTGA 2324
2701 CAGCCACTGGTAACAGATATGACAGAGCGAGGTATGTAGCGGTGTCTACAGTTCCTGA 2760
2325 AGTGTGGCTTAACCTACGCTACACTAGAGGACAGTATTTGTAATCTCGCTCTGCTGA 2384
2761 AGTGTGGCTTAACCTACGCTACACTAGAGGACAGTATTTGTAATCTCGCTCTGCTGA 2820
2385 AGCCAGTTACCTTCGGAAAAAGTGTGAGTCTTTGATCCGGCAAAACCAACCGCTG 2444
2821 AGCCAGTTACCTTCGGAAAAAGTGTGAGTCTTTGATCCGGCAAAACCAACCGCTG 2880
2445 GTAGCGGTGGTTTTTTTGTGTAAGCAGAGATTAACGCGAGAAAGGATCTCAAG 2504
2881 GTAGCGGTGGTTTTTTTGTGTAAGCAGAGATTAACGCGAGAAAGGATCTCAAG 2940
2505 AAGATCCTTTGATCTTTTCTACGGGCTGTGACGCTCAGTGGGACGAAACTCAGCTTAAG 2564
2941 AAGATCCTTTGATCTTTTCTACGGGCTGTGACGCTCAGTGGGACGAAACTCAGCTTAAG 3000

2565 GGATTTTGGTCATGAGATTTATCAAAAAGGATCTTCACTAGATCTCTTTTAAATTAATAAT 2624
3001 GGATTTTGGTCATGAGATTTATCAAAAAGGATCTTCACTAGATCTCTTTTAAATTAATAAT 3060
2625 GAAGTTTTAAATCAATCTAAAGTATATAGTAAACTTTGGTCTGACAGTTACCAATGCT 2684
3061 GAAGTTTTAAATCAATCTAAAGTATATAGTAAACTTTGGTCTGACAGTTACCAATGCT 3120
2685 TAATCAGTGAGCAGCCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGAC 2744
3121 TAATCAGTGAGCAGCCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGAC 3180
2745 TCCCGCTCGTGTAGATAACTACGATACGCGAGGCTTACATCTGCGCCAGTGTGCA 2804
3181 TCCCGCTCGTGTAGATAACTACGATACGCGAGGCTTACATCTGCGCCAGTGTGCA 3240
2805 TGATACCGCGAGACCCACGCTCACCGCTCCAGATTTATCAGCAATAAACACGAGCAGCG 2864
3241 TGATACCGCGAGACCCACGCTCACCGCTCCAGATTTATCAGCAATAAACACGAGCAGCG 3300
2865 GAAGGCGCGAGCGCAGAGTGTCTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAAT 2924
3301 GAAGGCGCGAGCGCAGAGTGTCTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAAT 3360
2925 GTTTCGCGGAGCTAGAGTAAGTTCGCCAGTTAAATAGTTTGCACAAGTGTGTCGA 2984
3361 GTTTCGCGGAGCTAGAGTAAGTTCGCCAGTTAAATAGTTTGCACAAGTGTGTCGA 3420
2985 TTGCTACAGGCACTGTGGTGTCAAGCTCGTCTGTTTGGTATGCTTTCATTCAGCTCCGTT 3044
3421 TTGCTACAGGCACTGTGGTGTCAAGCTCGTCTGTTTGGTATGCTTTCATTCAGCTCCGTT 3480
3045 CCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAGCGTTAGTCTCT 3104
3481 CCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAGCGTTAGTCTCT 3540
3105 TCGTCTCCGATCGTTGTGACAGTAAGTTCGCCAGCTGTTATCAGTATGCTGTTATGG 3164
3541 TCGTCTCCGATCGTTGTGACAGTAAGTTCGCCAGCTGTTATCAGTATGCTGTTATGG 3600
3165 CAGCACTGCATTAATCTCTTACTGTGTCATGCCATCGTAAGATGCTTTCTGTGACTGGTG 3224
3601 CAGCACTGCATTAATCTCTTACTGTGTCATGCCATCGTAAGATGCTTTCTGTGACTGGTG 3660
3225 AGTACTCAACCAAGTCATTTCTGAGAAATAGTATGCGCGACCGAGTTCCTCTGCCCCG 3284
3661 AGTACTCAACCAAGTCATTTCTGAGAAATAGTATGCGCGACCGAGTTCCTCTGCCCCG 3720
3285 CGTCAATACGGGATTAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGA 3344
3721 CGTCAATACGGGATTAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGA 3780
3345 AACGTTCTTCCGGGCGAAACTCTCAAGGATCTTACCGCTTTGAGATCCAGTTCGATGT 3404
3781 AACGTTCTTCCGGGCGAAACTCTCAAGGATCTTACCGCTTTGAGATCCAGTTCGATGT 3840
3405 AACCCACTCGTGCACCCAACTGATCTTCAAGCATCTTTTACTTTTCAACGAGTTTCTGGGT 3464
3841 AACCCACTCGTGCACCCAACTGATCTTCAAGCATCTTTTACTTTTCAACGAGTTTCTGGGT 3900
3465 GAGCAAAAACAGGAGGCAAAATGCGCAAAAAGGAAATAAGGCGACACGGAATGTT 3524
3901 GAGCAAAAACAGGAGGCAAAATGCGCAAAAAGGAAATAAGGCGACACGGAATGTT 3960
3525 GAATCTCATCTCTCTCTTTTCAATTTATTTAGAGCACTTTATCAGGCTTATGCTCA 3584
3961 GAATCTCATCTCTCTCTTTTCAATTTATTTAGAGCACTTTATCAGGCTTATGCTCA 4020
3585 TGAGCGGATACATATTTGAATGTTTGAATAAATAAAGGAAATAAAGGCGGTTCCGCGCACAT 3644
4021 TGAGCGGATACATATTTGAATGTTTGAATAAATAAAGGAAATAAAGGCGGTTCCGCGCACAT 4080
3645 TTCCCGGAAAAAGTGCCAC 3662

Db 1561 CACATTACTTTATATATATATTAATTTCAITTTATGTGATAATCGGAATCGCTTAAGAAAAA 1620
QY 1392 ----- 1391
Db 1621 AAGAGTCAATCCGTAGCGGAAAAAATAAATAATGAATAATCATTACCGAGGCATAAANAATA 1680
QY 1392 ----- 1391
Db 1681 TAGAGTGTACTAGAGGAGGCCAAGAGTAATAGAAAAAAGAAAAATTGCGGAAAGAGCTGTG 1740
QY 1392 ----- 1391
Db 1741 TTATGACTTCCCTGACTAATGCCGTGTTCAACGATACCTGGCAGTGACTCCTTAGCGCTC 1800
QY 1392 -----AATGACAGCGGTGCGCGCTGGGATATACGTACAGCGAGGACGGGTA 1438
Db 1801 ACCAAGCTCTTAAAAACGCGCGCTGCGAGAAATCGATATCAAGTTATCGATACCGTCG 1860
QY 1439 TCCTGGCTGGATCGCGAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGGC 1498
Db 1861 ACCTGAGGGGGGGCGGTACCCAGCTTTTGTTCCTTTTAGTGAGGGTTAATTGGCGCG 1920
QY 1499 TTGCGGTAAATCATGGTTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTTCCA 1558
Db 1921 TTGCGGTAAATCATGGTTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTTCCA 1980
QY 1559 CACAACATACGAGCGCGGAACATAAAGTGTAAAGCCTTGGGGTGCTTAATGAGTGAGCTAA 1618
Db 1981 CACAACATACGAGCGCGGAACATAAAGTGTAAAGCCTTGGGGTGCTTAATGAGTGAGCTAA 2040
QY 1619 CTCACATTAATTCGGTTGCGCTCACTGCGCGCTTTTCAGTFCGGGAAACCTGTCGCGCAG 1678
Db 2041 CTCACATTAATTCGGTTGCGCTCACTGCGCGCTTTTCAGTFCGGGAAACCTGTCGCGCAG 2100
QY 1679 CTCGATTAATGATCCGCCAAACCGCGGGAGAGGGGTTTGGTATTTGGCGCTCTTCC 1738
Db 2101 CTCGATTAATGATCCGCCAAACCGCGGGAGAGGGGTTTGGTATTTGGCGCTCTTCC 2160
QY 1739 GCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTATCAGCT 1798
Db 2161 GCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTATCAGCT 2220
QY 1799 CACTAAAGCGGTAAATACGGTTTATCCAGAAATCAGGGGATTAACCGAGAAAGAACATG 1858
Db 2221 CACTAAAGCGGTAAATACGGTTTATCCAGAAATCAGGGGATTAACCGAGAAAGAACATG 2280
QY 1859 TGAGCAAAAGGCCAGCAAAAGGCCAGAACCGTAAAGAACCGCGTTGCTGGCGTTTTC 1918
Db 2281 TGAGCAAAAGGCCAGCAAAAGGCCAGAACCGTAAAGAACCGCGTTGCTGGCGTTTTC 2340
QY 1919 CATAGGCTCGCGCCCTCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGA 1978
Db 2341 CATAGGCTCGCGCCCTCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGA 2400
QY 1979 AACCCGACAGGACTATAAAGATACAGGGGTTTCCCGCTGGAAAGTCCCTCGTGGCGCTCT 2038
Db 2401 AACCCGACAGGACTATAAAGATACAGGGGTTTCCCGCTGGAAAGTCCCTCGTGGCGCTCT 2460
QY 2039 CCGTTTCCGACCTCGCGCTTACGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTG 2098
Db 2461 CCGTTTCCGACCTCGCGCTTACGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTG 2520
QY 2099 GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGCTCCAAG 2158
Db 2521 GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGCTCCAAG 2580
QY 2159 CTGGGCTGTGTGACGAACCCCGTTTCAGCCCGACCGCTGCGCTTATCCCGTAACTAT 2218
Db 2581 CTGGGCTGTGTGACGAACCCCGTTTCAGCCCGACCGCTGCGCTTATCCCGTAACTAT 2640
QY 2219 CGTCTTGTGTCGAACCCCGTTAAGACGAGCTTATCGCCTAGTGCAGGCACTGGTAAC 2278

Db 2641 CGTCTTGAGTCCAAACCGGTAAAGACAGCACTTATCGCACTGCGCAGCAGCACTGGTAAC 2700
QY 2279 AGGATTAGCAGAGCGAGGTATGTAGCGGTGTCTACAGAGTTCTTGAAGTGGTGGCTTAAC 2338
Db 2701 AGGATTAGCAGAGCGAGGTATGTAGCGGTGTCTACAGACTTCTTGAAGTGGTGGCTTAAC 2760
QY 2339 TAGGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAGCCAGTACCTTC 2398
Db 2761 TAGGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAGCCAGTACCTTC 2820
QY 2399 GGAAGAAAGAGTTGGTAGCTCTTGATCCGCAAAACCAACCGCTGCTAGCGGTGGTTTTT 2458
Db 2821 GGAAGAAAGAGTTGGTAGCTCTTGATCCGCAAAAGAAACCAACCGCTGCTAGCGGTGGTTTTT 2880
QY 2459 TTTGTTTCAAGCAGCAGATTAACGCGCAAAAAAGAGTCTCAAGAAAGATCTTTTGATC 2518
Db 2881 TTTGTTTCAAGCAGCAGATTAACGCGCAAAAAAGAGTCTCAAGAAAGATCTTTTGATC 2940
QY 2519 TTTTCTACGGGTCTGACGCTCAGTGGAAAGAACTCAGTTTAAGGGAATTTTGGTCATG 2578
Db 2941 TTTTCTACGGGTCTGACGCTCAGTGGAAAGAACTCAGTTTAAGGGAATTTTGGTCATG 3000
QY 2579 AGATTATCAAAAGAGTCTTCACTAGATCTTTTAAATTTAAAAATGAAGTTTAAATCA 2638
Db 3001 AGATTATCAAAAGAGTCTTCACTAGATCTTTTAAATTTAAAAATGAAGTTTAAATCA 3060
QY 2639 ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA 2698
Db 3061 ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA 3120
QY 2699 CCTATCTCAGCGATCTGTCTATTTGTTTCCATCATAGTTGCTGACTCCCGCTCGTGTAG 2758
Db 3121 CCTATCTCAGCGATCTGTCTATTTGTTTCCATCATAGTTGCTGACTCCCGCTCGTGTAG 3180
QY 2759 ATAACTACCATACGGGAGGCTTTACCATCTGGCCCCAGTGTGCAATGATACCGCAGAC 2818
Db 3181 ATAACTACCATACGGGAGGCTTTACCATCTGGCCCCAGTGTGCAATGATACCGCAGAC 3240
QY 2819 CCACGCTCACCGCTCCAGATTTATCAGCAATAAACCCAGCCAGCCAGGAGGCGCAGCGC 2878
Db 3241 CCACGCTCACCGCTCCAGATTTATCAGCAATAAACCCAGCCAGCCAGGAGGCGCAGCGC 3300
QY 2879 AGAAGTGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCGCGGAGCT 2938
Db 3301 AGAAGTGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCGCGGAGCT 3360
QY 2939 AGAGTAAGTAGTTTCGCGCAGTTAATAGTTTGGCAAGTTGTTGCCATTTGCTACAGGCATC 2998
Db 3361 AGAGTAAGTAGTTTCGCGCAGTTAATAGTTTGGCAAGTTGTTGCCATTTGCTACAGGCATC 3420
QY 2999 GTGGTGTACGCTCGTCTGTTGGTATGGCTTCATTTCAGTCTCCGTTCCCAACGATCAAGG 3058
Db 3421 GTGGTGTACGCTCGTCTGTTGGTATGGCTTCATTTCAGTCTCCGTTCCCAACGATCAAGG 3480
QY 3059 CGAGTTACATGATCCCCCATGTTGCAAAAAAGCGGTTAGCTCTTCGTTCCCTCCGATC 3118
Db 3481 CGAGTTACATGATCCCCCATGTTGCAAAAAAGCGGTTAGCTCTTCGTTCCCTCCGATC 3540
QY 3119 GTTGTGAGAAAGTAAGTTGGCCGAGTGTATCACTCATGTTATGCGCAGCACTGCATAAT 3178
Db 3541 GTTGTGAGAAAGTAAGTTGGCCGAGTGTATCACTCATGTTATGCGCAGCACTGCATAAT 3600
QY 3179 TCTCTTACTGTCTATGCTCATCCGATCCGATGTTTCTGTGACTGGTGGTGAAGTCAACCAAG 3238
Db 3601 TCTCTTACTGTCTATGCTCATCCGATCCGATGTTTCTGTGACTGGTGGTGAAGTCAACCAAG 3660
QY 3239 TCATTCTGAGAAATAGTGTATGCGGCAACCGAGTGTCTTTCGCGCGGTCAATACCGGAT 3298
Db 3661 TCATTCTGAGAAATAGTGTATGCGGCAACCGAGTGTCTTTCGCGCGGTCAATACCGGAT 3720
QY 3299 AATACCGCCCATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAGAAAGTCTTTCGCGG 3358
Db 3721 AATACCGCCCATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAGAAAGTCTTTCGCGG 3780

```
QY 3359 CGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAAACCCACTCGTGCA 3418
Db
Db 3781 CGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAAACCCACTCGTGCA 3840
QY 3419 CCGAAGTATCTTACGATCTTTTACTTTCACAGCGTTTCTGGGTGAGCAAAAACAGGA 3478
Db
Db 3841 CCGAAGTATCTTACGATCTTTTACTTTCACAGCGTTTCTGGGTGAGCAAAAACAGGA 3900
QY 3479 AGGCAAAATGCCGCAAAAAGGGAATTAAGGCGACACGGAATGTTGAATCTCATATCTC 3538
Db
Db 3901 AGGCAAAATGCCGCAAAAAGGGAATTAAGGCGACACGGAATGTTGAATCTCATATCTC 3960
QY 3539 TTCCCTTTTCAATATTAATTAAGCACTTATCAGGCTTATGTTCTCATGAGCGGATACATA 3598
Db
Db 3961 TTCCCTTTTCAATATTAATTAAGCACTTATCAGGCTTATGTTCTCATGAGCGGATACATA 4020
QY 3599 TTGGAATGATTTAGAAAAATAAAACAAATAGGGGTTCCGCGACATTTCCCGCAAAAGTG 3658
Db
Db 4021 TTGGAATGATTTAGAAAAATAAAACAAATAGGGGTTCCGCGACATTTCCCGCAAAAGTG 4080
QY 3659 CCAC 3662
Db 4081 CCAC 4084

RESULT 9
US-09-993-170-1
; Sequence 1, Application US/099931170
; Patent No. 6743969
; GENERAL INFORMATION:
; APPLICANT: Valent, Barbara
; APPLICANT: Bryant, Gregory
; TITLE OF INVENTION: Modification of a Plant Disease Resistance Gene
; TITLE OF INVENTION: Specificity and Method for Engineering Altered
; TITLE OF INVENTION: Specificity
; FILE REFERENCE: BB1385 US NA
; CURRENT APPLICATION NUMBER: US/09/993,170
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 60/248,335
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 4119
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Plaemid
US-09-993-170-1

Query Match 62.2%; Score 2279.2; DB 4; Length 4119;
Best Local Similarity 76.6%; Pred. No. 0;
Matches 3151; Conservative 0; Mismatches 508; Indels 457; Gaps 11;

QY 1 CTAATTTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60
Db
Db 4 CTAATTTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 63
QY 61 ATTTTAAACCAATAGGCGCAAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 120
Db
Db 64 ATTTTAAACCAATAGGCGCAAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 123
QY 121 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAGCTGAGCTC 180
Db
Db 124 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAGCTGAGCTC 183
QY 181 CAAAGTCAAAAGGCGCAAAACCGTCTATCAGGCGCATGGCCCACTACGTGAACCATCACC 240
Db
Db 184 CAAAGTCAAAAGGCGCAAAACCGTCTATCAGGCGCATGGCCCACTACGTGAACCATCACC 243
QY 241 CTAATCAAGTTTTTGGGGTTCAGGTGCGTAAAGCACTAAATTCGGAACCCCTAAAGGGAG 300
```

```
Db
QY 244 CTAATCAAGTTTTTTGGGTCGAGGTCGCTAAAGCACTAAATTCGAAACCCCTAAAGGGAG 303
QY 301 CCCCGCATTTAGAGCTTCACGGGAAAGCCGCGCAACCTGGCGAGAAAGGAAGGCAAGAA 360
Db
Db 304 CCCCGCATTTAGAGCTTCACGGGAAAGCCGCGCAACCTGGCGAGAAAGGAAGGCAAGAA 363
QY 361 AGCGAAAGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAAACCAAC 420
Db
Db 364 AGCGAAAGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAAACCAAC 423
QY 421 CACACCCCGCGCTTAATGCGCGCTACAGGGCGGCTCCCATTCGCCATTCAGGCTGCG 480
Db
Db 424 CACACCCCGCGCTTAATGCGCGCTACAGGGCGGCTCCCATTCGCCATTCAGGCTGCG 483
QY 481 CAACTGTTGGGAAGGCGCATCGGTGCGGCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
Db
Db 484 CAACTGTTGGGAAGGCGCATCGGTGCGGCTCTTCGCTATTACGCCAGCTGGCGAAAGG 543
QY 541 GGGATGTCCTCAAGGCGATTAAGTTGGGTAAAGCCAGGGTTTTCCCAAGTCAACGCTTG 600
Db
Db 544 GGGATGTCCTCAAGGCGATTAAGTTGGGTAAAGCCAGGGTTTTCCCAAGTCAACGCTTG 603
QY 601 TAAACGAGCGGCGAGTGA-----GCGCGCTCGTT 630
Db
Db 604 TAAACGAGCGGCGAGTGAATTTGTAATACGACTCACTATAGGCGCAATTTGGGTACCGGGCC 663
QY 631 CATTCAAGCTTTTGAACCCGCTGGAGGACGGGCGAGACTCGCGTGCATAATGCTTTTACAG 690
Db
Db 684 CCCCTCGAGGTGCTATCCGAGGTAGGTCTAGGGGCTGATCCTCACAATATTTTGTAA 723
QY 691 CGTGATGGAGCAGATGAAGTGTCTGCACACGCTCGAGAACACGCGAGCT-----738
Db
Db 724 ATTTCAAAAGTCAGGAGCATGAATTAATGTAATTAATAATATGCGGCCCAACTCTTAC 783
QY 739 -----AGATTAACCTAGAAAGATAATCATATGTCAGCTAGCTTTAAAGATAATCAT 790
Db
Db 784 CTTATATAAAATTTGGGATGATATATAATAAAAGTGGACCTAATTAATCCTGCATAATAAT 843
QY 791 GCGTAAATTCAGCATGGGATCTGTAATAGACTCACTA--TAGGGGAAATGGGTACC 848
Db
Db 844 GCAGATATTAACACTAGCAAAATATAATTCGATAATATTAATGCTTAATAACGAT 903
QY 849 GGGCCCCCTCGAGGTGCGAGGTATCGATAGCTTGTATGGAATTCCTGAGCGCCGGG 908
Db
Db 904 TAATAAACCAATAGTTTACATCTTCTAAAGCTTTTGAATAAAAGTCAAGCTGAATAA 963
QY 909 GGATCCACTAGTTCTAGAGCGGCGCCACCGGGTGGAGCTCCAGCTTTTGTCCCTTTA 968
Db
Db 964 TAAATAAGTTGGCGTTGTTATAAATCGACCGCTTTCGCGCTTTAATGGTTTAATTCGGA 1023
QY 969 GTGAG-----GGTTAATTAGATCCCATCGCTCAATTTTACGAGACTATCTTTCT 1018
Db
Db 1024 TAGAACAATTTTGTGTTAATTTCCAAACATACAACAATATTCACCTGACTGAAAATCG 1083
QY 1019 AGGGTAAATCTAGCTGCTCAGGATCATA-----1047
Db
Db 1084 ACAGTTTTTGTGCAATCAACATTAATAATTAATAAACTTCTGCACAATAAACA 1143
QY 1048 -----TCGTCGGCTCTTTTCCGCTCAGTCATCGCCCAAGCTGGCGCTATCTGG 1099
Db
Db 1144 TTATTTTGTGATCGATGCTTTTTTATTCATTTTTCACACCGCTGCGATTCGG 1203
QY 1100 CAT-----1102
Db
Db 1204 CTTCAACCAATTTGGGACCTTTTTCACACCGGTTTACGATTACAAATCCAAATCCAAACC 1263
QY 1103 -----CGGGAGGAAGAGCCGCTGCTTT-1127
Db
Db 1264 ATATCCAGGAGATTTAAAGGCGGCTTATTTGAACGCTTATTTCCAAATGTTTCAAGATT 1323
QY 1128 -----TCCCGGAGGTTGAAGGCGATGGAAGAGTTTGGCGGAGTACGCTGCTCTCA 1182
Db
Db 1324 GCGAGGCTCCGAAATTCGTGCGCGCTTAAAGAGTTGTGCGGAGCTCGCTCGTGGGCT 1383
```


1183 TTGACGTTGAGCGAAACGCA--CGTTTACCATGATGATTCGGGAAG----- 1227
1384 ATCACGCCGTTAAATAATGCAATCGGTATTTAGATTAATCTTTAAACTGACAGCACAG 1443
1228 ----- 1227
1444 ATATTCAAACTCGGGTTCAAAGAAATTTAAACGAATTTAAACAGGAATGTAAACAGGACG 1503
1228 -----GTGTGGCCATGACGCGCTTTAAACGGTGAACCTGTTCGTTTCAGGCCACCT 1275
1504 CGGACGAATTTCTTAACTGACGATGATAAATGTTTATACGTGGTCCGAGAGGAG 1563
1276 GGAATACAGTTGCTCGCGGCTTTTCCGGACACAGTTCCGGATGCTGACGCCGAAGCGCA 1335
1564 TTCATAAATTGGCGTATGCACTTATTTAAACGAAGAAATTTGTTATATGCGCTCTTCT 1623
1336 TCAGCAACCGCAACATACCGCGCA----- 1362
1624 TCAACAACCCGTAACAGAGGGAATTAATGCGCGTAACCAAGATACAGTTATATTAC 1683
1363 -----GCCGAACCTGCCGTGCGGTG 1383
1684 ATGAATGTCATATAATTTTAAAGAGTGGAAAGATTTATGTTACGATGGATGGGA 1743
1384 TGCAGATTAAATGACAGCGGTGGCGCTGGGATATTACGTACAGGAGGACGGGTATCCTG 1443
1744 TTTACAAAATTGGATAGTACAGAAATTTAAACCCGACAGTTATGCTATTTTGGCAC 1803
1444 GCTGGATCGGCAGAAATGACATGATACCCCGTGGTTACCCGCGG----- 1493
1804 AATGTGCACGTTATAAATATTTGTTAAATAACGTTAGTGTGGAAATGGAGGGAATCGCGT 1863
1494 ----- 1493
1864 CAATCTAACATAGAGGATCCACTAGTTCTAGAGCGGCGGCACCGCGGTGGAGTCCA 1923
1494 -----CGCGTTGGCGTAATCATGGTCATAGCTGT 1523
1924 GCTTTTGTTCCTTTAGTGGGGTTAAATTTGAGCTTGGCGTAATCATGGTCATAGCTGT 1983
1524 TTTCTGTGTGAATTTGTTATCCGTCACAAATTTCCACACATACAGCGCGGAGATAA 1583
1984 TTTCTGTGTGAATTTGTTATCCGCTCACAAATTTCCACACATACAGCGCGGAGATAA 2043
1584 AGTTAAAGCTGGGTGCTTAATGATGAGCTAACTCAATTAATTTGGTTGCGCTCAC 1643
2044 AGTTAAAGCTGGGTGCTTAATGATGAGCTAACTCAATTAATTTGGTTGCGCTCAC 2103
1644 TGCCTGCTTTCCAGTCGGGAAACCTGCTGTCGAGCTGCAATTAATGAATCGGCCAACGCG 1703
2104 TGCCTGCTTTCCAGTCGGGAAACCTGCTGTCGAGCTGCAATTAATGAATCGGCCAACGCG 2163
1704 CGGGAGAGCGGTTTGGTATTTGGCGCTCTTCCGCTTCTCGCTCAGCTGCTGCTG 1763
2164 CGGGAGAGCGGTTTGGTATTTGGCGCTCTTCCGCTTCTCGCTCAGCTGCTGCTG 2223
1764 GCTCGGTGCTTGGCTGCGCGAGCGGTATCAGCTCACTCAAGGCGGTAAATACGGTTAT 1823
2224 GCTCGGTGCTTGGCTGCGCGAGCGGTATCAGCTCACTCAAGGCGGTAAATACGGTTAT 2283
1824 CCACAGAATCAGGGGATAACGAGGAAAGAAACATGTGAGCAAAAGCCAGCAAAAGGCCA 1883
2284 CCACAGAATCAGGGGATAACGAGGAAAGAAACATGTGAGCAAAAGCCAGCAAAAGGCCA 2343
1884 GGAAACGTTAAAGCGCGGTTGCTGGCGTTTTCATAGGCTTCCGCCCTTCCGAGCAG 1943
2344 GGAAACGTTAAAGCGCGGTTGCTGGCGTTTTCATAGGCTTCCGCCCTTCCGAGCAG 2403
1944 ATCAAAAAATCGACCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATTAAGATACC 2003
2404 ATCAAAAAATCGACCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATTAAGATACC 2463

2004 AGCGGTTTCCCGCTGGAAGCTCCCTGCTGGCTCTCTGTTCCGACCCCTGCCGCTTACCG 2063
2464 AGCGGTTTCCCGCTGGAAGCTCCCTGCTGGCTCTCTGTTCCGACCCCTGCCGCTTACCG 2523
2064 GATACCTGTCGCGCTTTCTCCCTTCCGGAAGCGTGGCGCTTTCTCATAGCTCAACGCTGA 2123
2524 GATACCTGTCGCGCTTTCTCCCTTCCGGAAGCGTGGCGCTTTCTCATAGCTCAACGCTGA 2583
2124 GGTATCTCAGTTCCGTTGAGTTCGCTTCCGCTCCTCAAGCTGGGCTGTGTGACGAAACCCCG 2183
2584 GGTATCTCAGTTCCGTTGAGTTCGCTTCCGCTCCTCAAGCTGGGCTGTGTGACGAAACCCCG 2643
2184 TTCAGCCCGACCGCTGCGCTTATCCGCTAACTATGCTCTTGGAGTCCAAACCCGCTGAAC 2243
2644 TTCAGCCCGACCGCTGCGCTTATCCGCTAACTATGCTCTTGGAGTCCAAACCCGCTGAAC 2703
2244 AGGACTTATCGCACCTGCGAGCAGCCACCTGTAACAGGATTAAGCAGAGGAGTATGTAG 2303
2704 AGGACTTATCGCACCTGCGAGCAGCCACCTGTAACAGGATTAAGCAGAGGAGTATGTAG 2763
2304 GCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTACCGCTACACTAGAAGGACAGTAT 2363
2764 GCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTACCGCTACACTAGAAGGACAGTAT 2823
2364 TTTGGTATCTGCGCTGCTGTAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGAT 2423
2824 TTTGGTATCTGCGCTGCTGTAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGAT 2883
2424 CCGGCAAAACAAACACCGCTGCTGAGCGGTGTTTTTTGTTTTCGAAAGCAGAGATTACGC 2483
2884 CCGGCAAAACAAACACCGCTGCTGAGCGGTGTTTTTTGTTTTCGAAAGCAGAGATTACGC 2943
2484 GCAGAAAAAAGAGTCTCAAGAGAGTCTTTGATCTTTTCTACGGGCTCTGACGCTCAGT 2543
2944 GCAGAAAAAAGAGTCTCAAGAGAGTCTTTGATCTTTTCTACGGGCTCTGACGCTCAGT 3003
2544 GGAAAGAAATCTCAGTTAAGGATTTTGGTCAATGAGATTAACAAGAGTCTTCACT 2603
3004 GGAAAGAAATCTCAGTTAAGGATTTTGGTCAATGAGATTAACAAGAGTCTTCACT 3063
2604 AGATCTTTTAAATTAATAATGAATTTTAAATCAATCTAAAGATATATAGTAAACTT 2663
3064 AGATCTTTTAAATTAATAATGAATTTTAAATCAATCTAAAGATATATAGTAAACTT 3123
2664 GGTCTGACAGTTTACCAATGCTTAACTCAGTGGGACCTTATCTCAGCGATCTGTCTATTTC 2723
3124 GGTCTGACAGTTTACCAATGCTTAACTCAGTGGGACCTTATCTCAGCGATCTGTCTATTTC 3183
2724 GTTCTCATAGTTGCTGACTCCCGCTGCTGTAGATAACTAGATACGGGAGGCTTAC 2783
3184 GTTCTCATAGTTGCTGACTCCCGCTGCTGTAGATAACTAGATACGGGAGGCTTAC 3243
2784 CATCTGGCCCCAGTGTCAATGATACCGGAGACCCAGCTCACCGGCTCCAGATTTAT 2843
3244 CATCTGGCCCCAGTGTCAATGATACCGGAGACCCAGCTCACCGGCTCCAGATTTAT 3303
2844 CAGCAATAAACACAGCCAGCCGGAAGGCGCAGAGAGTGGTCTCTGCAACTTTATCCG 2903
3304 CAGCAATAAACACAGCCAGCCGGAAGGCGCAGAGAGTGGTCTCTGCAACTTTATCCG 3363
2904 CTTCCATCCAGTCTATTAATTTGTCGGGAAAGCTAGAGTAAGTATGCTGCGCAATTAATA 2963
3364 CTTCCATCCAGTCTATTAATTTGTCGGGAAAGCTAGAGTAAGTATGCTGCGCAATTAATA 3423
2964 GTTTCGGCAACCTGTTGCTCAATTTGCTACAGGATCGTGGTGTACGCTCGTTGGTA 3023
3424 GTTTCGGCAACCTGTTGCTCAATTTGCTACAGGATCGTGGTGTACGCTCGTTGGTA 3483
3024 TGGCTTCAATTCAGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATCCCCCATGTTGT 3083
3484 TGGCTTCAATTCAGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATCCCCCATGTTGT 3543
3084 GCAAAAAAGCGGTAGCTCCTTCGGTCTCCGATCGTTGTCAGAAAGTAAGTTGGCGCGCAG 3143

Db	3544	GCAGAAAAAGCGGTAGTCTCTTCGGTCTCCGATCGTTCAGAAAGTAAGTTGGCGCAG	3603	
Qy	3144	TGTTATCACTCATGGTTATGGCAGACACTGCATAATCTCTTACTGTCTGATGCCATCGGTAA	3203	
Db	3604	TGTTATCACTCATGGTTATGGCAGACACTGCATAATCTCTTACTGTCTGATGCCATCGGTAA	3663	
Qy	3204	GATGCTTTTCTGTGACTGGTGAAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGC	3263	
Db	3664	GATGCTTTTCTGTGACTGGTGAAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGC	3723	
Qy	3264	GACCGAGTTGCTCTTGGCCCGGCTCAATACGGGATAATACCGCGCCACATACGAAACTT	3323	
Db	3724	GACCGAGTTGCTCTTGGCCCGGCTCAATACGGGATAATACCGCGCCACATACGAAACTT	3783	
Qy	3324	TAAAGTGTCTATCTGAGAAACGTTCTTCGGGGGGAACCTCTCAAGGATCTTACCGC	3383	
Db	3784	TAAAGTGTCTATCTGAGAAACGTTCTTCGGGGGGAACCTCTCAAGGATCTTACCGC	3843	
Qy	3384	TGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCCACTCTTCAGCATCTTTTA	3443	
Db	3844	TGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCCACTCTTCAGCATCTTTTA	3903	
Qy	3444	CTTTCAACGCTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCGCAAAAAAGGGAA	3503	
Db	3904	CTTTCAACGCTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCGCAAAAAAGGGAA	3963	
Qy	3504	TAAAGGCGACCGGAATGTTCAATACATCATCTCTCTCTTTTCAATATTTAGAACCA	3563	
Db	3964	TAAAGGCGACCGGAATGTTCAATACATCATCTCTCTCTTTTCAATATTTAGAACCA	4023	
Qy	3564	TTTTATCAGGGTTATGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAC	3623	
Db	4024	TTTTATCAGGGTTATGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAC	4083	
Qy	3624	AAATAGGGTTCCGCGCACATTTCCCGGAAAAAGTGC	3659	
Db	4084	AAATAGGGTTCCGCGCACATTTCCCGGAAAAAGTGC	4119	
RESULT 10				
US-09-138-024-20				
; Sequence 20, Application US/09138024A				
; Patent No. 6004779				
; GENERAL INFORMATION:				
; APPLICANT: Bradley, John D.				
; APPLICANT: Thompson, Craig M.				
; APPLICANT: Moore, Jeffrey B.				
; APPLICANT: Wobbe, C. Richard				
; APPLICANT: Healy, Judith M.				
; APPLICANT: Donnelly, Caroline E.				
; TITLE OF INVENTION: REGULATED GENE EXPRESSION IN YEAST				
; FILE REFERENCE: 0342/1D46SUS1				
; CURRENT APPLICATION NUMBER: US/09/138,024A				
; CURRENT FILING DATE: 1998-08-21				
; EARLIER APPLICATION NUMBER: 60/056,719				
; EARLIER FILING DATE: 1997-08-22				
; NUMBER OF SEQ ID NOS: 24				
; SOFTWARE: FastSeq for Windows Version 3.0				
; SEQ ID NO 20				
; LENGTH: 7102				
; TYPE: DNA				
; ORGANISM: Artificial Sequence				
; FEATURE:				
; OTHER INFORMATION: Plasmid p2M195				
US-09-138-024-20				
Query Match 59.3%; Score 2172.8; DB 3; Length 7102;				
Best Local Similarity 99.7%; Pred. No. 0;				
Matches 217; Conservative 7; Mismatches 0; Indels 0; Gaps 0;				
Qy	1479	GAGTTACCGCGCGCGCGCTTGGCGTAAATCATGTCATAGCTGTCTCTGTGTAATTT	1538	

Db	3341	GAGTCAGCGGCATGCAAGCTTGGCGTAAATCATGTCATAGCTGTTCTCTGTGTGAAATTT	3400	
Qy	1539	GTTATCCGCTCAATTTCCACAACATACGAGCGGAAGCATAAAGTGTAAAGCTCGG	1598	
Db	3401	GTTATCCGCTCAATTTCCACAACATACGAGCGGAAGCATAAAGTGTAAAGCTCGG	3460	
Qy	1599	GTGCTTAATGAGTGAGCTAACTCACATTAATTTGCGTTGCGCTCACTGCGCGCTTCCAGT	1658	
Db	3461	GTGCTTAATGAGTGAGCTAACTCACATTAATTTGCGTTGCGCTCACTGCGCGCTTCCAGT	3520	
Qy	1659	CGGAAAACTGTGTCGCGCAGCTGCAATTAATGATCGGCCAACCGCGGGGAGAGCGGTT	1718	
Db	3521	CGGAAAACTGTGTCGCGCAGCTGCAATTAATGATCGGCCAACCGCGGGGAGAGCGGTT	3580	
Qy	1719	TGCGTATTTGGGCGCTCTTCGCTTCCCTCGCTCACTGACTCGCTGCGCTCGGTTCGCG	1778	
Db	3581	TGCGTATTTGGGCGCTCTTCGCTTCCCTCGCTCACTGACTCGCTGCGCTCGGTTCGCG	3640	
Qy	1779	TGCGCGAGCGGTATCAGTCACTCAAAAGGCGGTAAATACGGTTATCCACAGAAATCAGGG	1838	
Db	3641	TGCGCGAGCGGTATCAGTCACTCAAAAGGCGGTAAATACGGTTATCCACAGAAATCAGGG	3700	
Qy	1839	ATAACGAGGAAAGAAACATGTGAGCAAAAAGCCAGCAAAAAGCCAGGAACCGTAAAGG	1898	
Db	3701	ATAACGAGGAAAGAAACATGTGAGCAAAAAGCCAGCAAAAAGCCAGGAACCGTAAAGG	3760	
Qy	1899	CGCGTTCGTCGCGCTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGAC	1958	
Db	3761	CGCGTTCGTCGCGCTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGAC	3820	
Qy	1959	GCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACAGCGGTTTCCCGCTG	2018	
Db	3821	GCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACAGCGGTTTCCCGCTG	3880	
Qy	2019	GAAGTCCCTGTGCGGCTCTCTGTTCCGACCTGCGCGTTACCGGATACCTGTCGCGCT	2078	
Db	3881	GAAGTCCCTGTGCGGCTCTCTGTTCCGACCTGCGCGTTACCGGATACCTGTCGCGCT	3940	
Qy	2079	TTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG	2138	
Db	3941	TTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG	4000	
Qy	2139	TGTAGTCTGTGCTCCAAAGTGGGCTGTGTGCAGAAACCCCGCTTACGCCGACCGCT	2198	
Db	4001	TGTAGTCTGTGCTCCAAAGTGGGCTGTGTGCAGAAACCCCGCTTACGCCGACCGCT	4060	
Qy	2199	GGCCTTATCCGTTAACTATGCTTTGAGTCCAAACCCGCTAAGACACGACTTATCGCAC	2258	
Db	4061	GGCCTTATCCGTTAACTATGCTTTGAGTCCAAACCCGCTAAGACACGACTTATCGCAC	4120	
Qy	2259	TGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTCAGAGT	2318	
Db	4121	TGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTCAGAGT	4180	
Qy	2319	TCTTGAAGTGTGGCTTAACCTACCGCTACATAGAAAGACAGTATTGATGATCTCGGCTC	2378	
Db	4181	TCTTGAAGTGTGGCTTAACCTACCGCTACATAGAAAGACAGTATTGATGATCTCGGCTC	4240	
Qy	2379	TGCTGAAGCCAGTTACCTTCGGNAAAAGATTGGTAGCTCTTGATCCGCGCAACCAACCA	2438	
Db	4241	TGCTGAAGCCAGTTACCTTCGGNAAAAGATTGGTAGCTCTTGATCCGCGCAACCAACCA	4300	
Qy	2439	CCGCTGTAGCGGTGGTTTTTTTTTTTGTTCGAAGCAGCAGATTACCGCGCAGAAAAAGGAT	2498	
Db	4301	CCGCTGTAGCGGTGGTTTTTTTTTTTGTTCGAAGCAGCAGATTACCGCGCAGAAAAAGGAT	4360	
Qy	2499	CTCAAGAAAGATCTTTTGATCTTTTCTACGGGTCTGACGGTCTCAGTGGGAACGAACTCAC	2558	
Db	4361	CTCAAGAAAGATCTTTTGATCTTTTCTACGGGTCTGACGGTCTCAGTGGGAACGAACTCAC	4420	
Qy	2559	GTTAAGGATTTGGTCACTGAGATTATCAAAAAGGATCTTCACTAGATCTCTTTTAAATTT	2618	
Db	4421	GTTAAGGATTTGGTCACTGAGATTATCAAAAAGGATCTTCACTAGATCTCTTTTAAATTT	4480	

Qy=	2619	AAAAATGAAGTTTAAATCAATCTAAAGTATATATAGTAAAACTTGCTGTGACAGATTACC	2678
Db	4481	AAAAATGAAGTTTAAATCAATCTAAAGTATATATAGTAAAACTTGCTGTGACAGATTACC	4540
Qy	2679	AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGCTATTTCGTTTCATCCATAGTTG	2738
Db	4541	AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGCTATTTCGTTTCATCCATAGTTG	4600
Qy	2739	CCTGACTCCCCGTCGTGTAGATAAATCAGATACGGGAGGCTTACCACTTGGCCCCCAGTG	2798
Db	4601	CCTGACTCCCCGTCGTGTAGATAAATCAGATACGGGAGGCTTACCACTTGGCCCCCAGTG	4660
Qy	2799	CTGCAATGATACCGCGAGACCCACGCTCACGGCTCCAGATTTATCAGCAATAAACACGAC	2858
Db	4661	CTGCAATGATACCGCGAGACCCACGCTCACGGCTCCAGATTTATCAGCAATAAACACGAC	4720
Qy	2859	CAGCCGGAAGGGCCGACGCGAGAAGTGCTCTGTCAAATTTATCCGCTCCATCCAGTCTA	2918
Db	4721	CAGCCGGAAGGGCCGACGCGAGAAGTGCTCTGTCAAATTTATCCGCTCCATCCAGTCTA	4780
Qy	2919	TTAATTGTTCCGGGAAGCTAGAGTAAGTAGTTCCGCAGTTTAATAGTTTGGCAACGTTG	2978
Db	4781	TTAATTGTTCCGGGAAGCTAGAGTAAGTAGTTCCGCAGTTTAATAGTTTGGCAACGTTG	4840
Qy	2979	TTGCAATTGCTACAGGCATCGTGGTGTCACGCTCGTCTGTTGGTATAGCTTCATTCAGCT	3038
Db	4841	TTGCAATTGCTACAGGCATCGTGGTGTCACGCTCGTCTGTTGGTATAGCTTCATTCAGCT	4900
Qy	3039	CCGGTTCCTCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTA	3098
Db	4901	CCGGTTCCTCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTA	4960
Qy	3099	GCTCTTTCGGTCTCCGATCGTTCTCAGAAAGTAGTTGGCGCAGTGTATCACTCATGG	3158
Db	4961	GCTCTTTCGGTCTCCGATCGTTCTCAGAAAGTAGTTGGCGCAGTGTATCACTCATGG	5020
Qy	3159	TTATGGCAGCATGTCATAAATCTCTTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGA	3218
Db	5021	TTATGGCAGCATGTCATAAATCTCTTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGA	5080
Qy	3219	CTGGTGTAGTACTCAACCAAGTCAATCTCTGAGAAATAGTGTATGCGGCGACCGAGTGTCTTT	3278
Db	5081	CTGGTGTAGTACTCAACCAAGTCAATCTCTGAGAAATAGTGTATGCGGCGACCGAGTGTCTTT	5140
Qy	3279	GCCGGCGTCAATACGGGATAATACCGGCCACATACGAGAACCTTTAAAAGTGTCTATCA	3338
Db	5141	GCCGGCGTCAATACGGGATAATACCGGCCACATACGAGAACCTTTAAAAGTGTCTATCA	5200
Qy	3339	TTGGAAAAAGTCTTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTT	3398
Db	5201	TTGGAAAAAGTCTTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTT	5260
Qy	3399	CGATGTAACCCACTCGTGTGCAACCACTGATCTTCAGCATCTTTTTCATTTTCAACGCGTTT	3458
Db	5261	CGATGTAACCCACTCGTGTGCAACCACTGATCTTCAGCATCTTTTTCATTTTCAACGCGTTT	5320
Qy	3459	CTGGGTGAGCAAAAACAGGAGGCAAAATCCGCAAAAAGGGAATTAAGGCGCGACCGGA	3518
Db	5321	CTGGGTGAGCAAAAACAGGAGGCAAAATCCGCAAAAAGGGAATTAAGGCGCGACCGGA	5380
Qy	3519	AATGTTGAATCTCATACTCTTCTCTTTTCAAATATATGAAGCATTTATCAGGGTTATT	3578
Db	5381	AATGTTGAATCTCATACTCTTCTCTTTTCAAATATATGAAGCATTTATCAGGGTTATT	5440
Qy	3579	GTCTCATGAGCGGATACATATTTGAAATGTATTTAGAAAAATAAATAAATAGGGTTCCTCGC	3638
Db	5441	GTCTCATGAGCGGATACATATTTGAAATGTATTTAGAAAAATAAATAAATAGGGTTCCTCGC	5500
Qy	3639	GCACATTTCCCGAAAAAGTGCCAC	3692
Db	5501	GCACATTTCCCGAAAAAGTGCCAC	5524

```

RESULT 11
US-09-404-066-20
; Sequence 20, Application US/09404066
; Patent No. 6365409
; GENERAL INFORMATION:
; APPLICANT: Bradley, John D.
; APPLICANT: Thompson, Craig M.
; APPLICANT: Moore, Jeffrey B.
; APPLICANT: Wobbe, C. Richard
; APPLICANT: Healy, Judith M.
; APPLICANT: Donnelly, Caroline E.
; TITLE OF INVENTION: REGULATED GENE EXPRESSION IN YEAST
; FILE REFERENCE: 0342/1D469US1
; CURRENT APPLICATION NUMBER: US/09/404,066
; CURRENT FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/138,024
; PRIOR FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: 60/056,719
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 7102
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid pZM195
US-09-404-066-20

```

Query Match	59.3%	Score 2172.8	DB 3	Length 7102
Best Local Similarity	99.7%	Pred. No. 0		
Matches 2177	Conservative 0	Mismatches 7	Indels 0	Gaps 0
Qy	1479	GAGTTACCGGGCGCGCGCTTCGCGCTTAATCATGTGTCATAGCTGTATCGTGTCTCTGTGTGAATTT	1538	
Db	3341	GAGTCACCGGCATGCAAGCTTGGCGTATCATGTGTCATAGCTGTTCCTGTGTGAATTT	3400	
Qy	1539	GTTTATCCGCTCACAATTTCCACACAAATACGAGCCGGGAAGCATAAAGTGTAAAGCCCTGGG	1598	
Db	3401	GTTTATCCGCTCACAATTTCCACACAAATACGAGCCGGGAAGCATAAAGTGTAAAGCCCTGGG	3460	
Qy	1599	GTCCCTAATGAGTGAGCTAACTCACATTAATTCGTTGGCTTCGCTCACGTCCCGCTTTCACGT	1658	
Db	3461	GTCCCTAATGAGTGAGCTAACTCACATTAATTCGTTGGCTTCGCTCACGTCCCGCTTTCACGT	3520	
Qy	1659	CGGGAAACCTGTTCGTGTCAGCTCATTAAATGAATCGGCCAAACGCGGGGGAGAGGCGGTT	1718	
Db	3521	CGGGAAACCTGTTCGTGCCAGCTCATTAAATGAATCGGCCAAACGCGGGGGAGAGGCGGTT	3580	
Qy	1719	TGGGTATTTGGGCGCTCTTCGCGTTCCTCGCTCACTGACTCGCTGCGCTCGGTGTTTCGGC	1778	
Db	3581	TGGGTATTTGGGCGCTCTTCGCGTTCCTCGCTCACTGACTCGCTGCGCTCGGTGTTTCGGC	3640	
Qy	1779	TGCGGGGAGCGGTATCAGCTCAGCTCAAAGSCGTAAATACGTTTATCCACAGAAATCAGGGG	1838	
Db	3641	TGCGGGGAGCGGTATCAGCTCAGCTCAAAGSCGTAAATACGTTTATCCACAGAAATCAGGGG	3700	
Qy	1839	ATAACGCAGGAAAGAACTGTGAGCAAAAGGCGCAGCAAAAGGCCAGGAACCGTAAAAAAGG	1998	
Db	3701	ATAACGCAGGAAAGAACTGTGAGCAAAAGGCGCAGCAAAAGGCCAGGAACCGTAAAAAAGG	3760	
Qy	1899	CCGCGTTGCTGGGCGTTTTTCCATAGGCTCGGCCCGCCCTGACGAGCATCAAAAATTCGCAC	1958	
Db	3761	CCGCGTTGCTGGGCGTTTTTCCATAGGCTCGGCCCGCCCTGACGAGCATCAAAAATTCGCAC	3820	
Qy	1959	GCTCAAAGTCAGAGGTGGCGAAAACCCACAGGACTATAAAGATACACGAGCGGTTTCCCGCTG	2018	
Db	3821	GCTCAAAGTCAGAGGTGGCGAAAACCCACAGGACTATAAAGATACACGAGCGGTTTCCCGCTG	3880	
Qy	2019	GAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCTTCGCGCTTACCGGATACCTGTCCGCCT	2078	
Db	3881	GAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCTTCGCGCTTACCGGATACCTGTCCGCCT	3940	

QY 2079 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGG 2138
Db 3941 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGG 4000
QY 2139 TGTAGGTGTTCGCTCCAAAGCTGGGCTGTGTGCAAGAACCCCGCTTGACGCCGACCGCT 2198
Db 4001 TGTAGGTGTTCGCTCCAAAGCTGGGCTGTGTGCAAGAACCCCGCTTGACGCCGACCGCT 4060
QY 2199 GGGCTTTATCCGGTAACATCGTCTTGAGTCAACCCGGTAAGACACGACTTTATCGCCAC 2258
Db 4061 GGGCTTTATCCGGTAACATCGTCTTGAGTCCNACCCGGTAAGACACGACTTTATCGCCAC 4120
QY 2259 TGGCAGCAGCCACTCGGTAA CAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT 2318
Db 4121 TGGCAGCAGCCACTCGGTAA CAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT 4180
QY 2319 TCTTGAAGTGTGGCTTAACATACGGGTACATAGAGGACAGTATTTGGTATCTCGGCTC 2378
Db 4181 TCTTGAAGTGTGGCTTAACATACGGGTACATAGAGGACAGTATTTGGTATCTCGGCTC 4240
QY 2379 TGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGTAGCTCTTGATCCGCGCAACAAACA 2438
Db 4241 TGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGTAGCTCTTGATCCGCGCAACAAACA 4300
QY 2439 CGCTGGTAGCGGTGGTTTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAGGAT 2498
Db 4301 CGCTGGTAGCGGTGGTTTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAGGAT 4360
QY 2499 CTCAGAAGATCTTTTGTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAAGAACTCAC 2558
Db 4361 CTCAGAAGATCTTTTGTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAAGAACTCAC 4420
QY 2559 GTTAAGGATTTTGGTCTAGATGATTATCAAAAGGATCTTCACTAGATCTCTTTTAAAT 2618
Db 4421 GTTAAGGATTTTGGTCTAGATGATTATCAAAAGGATCTTCACTAGATCTCTTTTAAAT 4480
QY 2619 AAAATGAAGTTTTAAATCAATTAAGTATATATGAGTAACTTGGTCTGACAGTTACC 2678
Db 4481 AAAATGAAGTTTTAAATCAATTAAGTATATATGAGTAACTTGGTCTGACAGTTACC 4540
QY 2679 AATGCTTAATCAGTGAGGACCTATCTCAGCGATCTGTCTATTTCTGTTTCATCCATAGTTG 2738
Db 4541 AATGCTTAATCAGTGAGGACCTATCTCAGCGATCTGTCTATTTCTGTTTCATCCATAGTTG 4600
QY 2739 CCTGACTCCCGTGTGTAGATAACTACGATACGGAGGGCTTACCATCTGSCCCAGTG 2798
Db 4601 CCTGACTCCCGTGTGTAGATAACTACGATACGGAGGGCTTACCATCTGSCCCAGTG 4660
QY 2799 CTGCAATGATACCGGAGACCCACGCTCACCGCTCCAGATTATCAGCAATAAACCCAGC 2858
Db 4661 CTGCAATGATACCGGAGACCCACGCTCACCGCTCCAGATTATCAGCAATAAACCCAGC 4720
QY 2859 CAGCCGAAGGCGCGAGCGAGAGTGTCTGTGCAACTTTTATCCGCTCCATCCAGTCTA 2918
Db 4721 CAGCCGAAGGCGCGAGCGAGAGTGTCTGTGCAACTTTTATCCGCTCCATCCAGTCTA 4780
QY 2919 TTAATTTGTCGGGAAGCTAGAGTAAGTAGTTCCGCAAGTTAATAGTTTGGCAAGTTG 2978
Db 4781 TTAATTTGTCGGGAAGCTAGAGTAAGTAGTTCCGCAAGTTAATAGTTTGGCAAGTTG 4840
QY 2979 TTGCCATTGCTACGGCATCGTGTGTCACTCGTCTGTTGGTATGGCTTCATTCACT 3038
Db 4841 TTGCCATTGCTACGGCATCGTGTGTCACTCGTCTGTTGGTATGGCTTCATTCACT 4900
QY 3039 CCGGTTCCCAAGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTA 3098
Db 4901 CCGGTTCCCAAGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTA 4960
QY 3099 GCTCCCTCGGCTCCGATCGTTGTCAGAGTAAGTTGGCGCAGTGTATCACTCATTG 3158
Db 4961 GCTCCCTCGGCTCCGATCGTTGTCAGAGTAAGTTGGCGCAGTGTATCACTCATG 5020

QY 3159 TTATGGCAGCACTGCATAAATTTCTTTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGA 3218
Db 5021 TTATGGCAGCACTGCATAAATTTCTTTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGA 5080
QY 3219 CTGGTGAGTACTCAACCAAGTCATTTCTGAGATAGTGTATGCGGCGACCGAGTTGCTCTT 3278
Db 5081 CTGGTGAGTACTCAACCAAGTCATTTCTGAGATAGTGTATGCGGCGACCGAGTTGCTCTT 5140
QY 3279 GCCCGGCTCAATACGGGATAATACCGGCCACATAGCAGAACTTTTAAAGTGTCTCATCA 3338
Db 5141 GCCCGGCTCAATACGGGATAATACCGGCCACATAGCAGAACTTTTAAAGTGTCTCATCA 5200
QY 3339 TTGGAAGAAAGTTCTTTCGGGGGAAACTCTCAAGATCTTACCGCTGTGTAGATCCAGTT 3398
Db 5201 TTGGAAGAAAGTTCTTTCGGGGGAAACTCTCAAGATCTTACCGCTGTGTAGATCCAGTT 5260
QY 3399 CGATGTAAACCACTCGTCGACCCCACTGATCTTTCAGCATCTTTTACTTTTACCGGTTT 3458
Db 5261 CGATGTAAACCACTCGTCGACCCCACTGATCTTTCAGCATCTTTTACTTTTACCGGTTT 5320
QY 3459 CTGGTGAGCAAAAACAGGAAGGCAAAATGCGCAAAAAAGGAATAAGGGCGACACGGA 3518
Db 5321 CTGGTGAGCAAAAACAGGAAGGCAAAATGCGCAAAAAAGGAATAAGGGCGACACGGA 5380
QY 3519 AATGTTGAATATCTCATCTCTTCTTTTCAATATTTGAAGCATTTATCAGGTTTATT 3578
Db 5381 AATGTTGAATATCTCATCTCTTCTTTTCAATATTTGAAGCATTTATCAGGTTTATT 5440
QY 3579 GTCTCATGAGCGGATACATATTTGATGATTTTGAATGATTTAGAAAAATAAACAATAGGGTTCCGC 3638
Db 5441 GTCTCATGAGCGGATACATATTTGATGATTTTGAATGATTTAGAAAAATAAACAATAGGGTTCCGC 5500
QY 3639 GCACATTTTCCCGAAAAGTGCCAC 3662
Db 5501 GCACATTTTCCCGAAAAGTGCCAC 5524

RESULT 12
US-09-573-322-20
; Sequence 20, Application US/09573322
; Patent No. 6531289
; GENERAL INFORMATION:
; APPLICANT: Bradley, John D.
; APPLICANT: Thompson, Craig M.
; APPLICANT: Moore, Jeffrey B.
; APPLICANT: Wobbe, C. Richard
; APPLICANT: Bailey, David A.
; TITLE OF INVENTION: Regulated Gene Expression in Yeast and
; FILE REFERENCE: 0342/1D469-US4
; CURRENT APPLICATION NUMBER: US/09/573,322
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/404,066
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/138,024
; PRIOR FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: 60/056,719
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 7102
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ZM195 plasmid
US-09-573-322-20

Query Match 59.3%; Score 2172.8; DB 4; Length 7102;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1479 GAGTTACCCGGCGGCGCTTGGCGTAATCATGTGCTAGTCTGTCTGTGAAAT 1538

RESULT 13

US-09-138-024-21
; Sequence 21, Application US/09138024A
; Patent No. 6004779
; GENERAL INFORMATION:
; APPLICANT: Bradley, John D.
; APPLICANT: Thompson, Craig M.
; APPLICANT: Moore, Jeffrey B.
; APPLICANT: Wobbe, C. Richard
; APPLICANT: Healy, Judith M.
; APPLICANT: Donnelly, Caroline E.
; TITLE OF INVENTION: REGULATED GENE EXPRESSION IN YEAST
; FILE REFERENCE: 0342/1D469US1
; CURRENT APPLICATION NUMBER: US/09/138,024A
; CURRENT FILING DATE: 1998-08-21
; EARLIER APPLICATION NUMBER: 60/056,719
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FaestSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 7333
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid pZM197
US-09-138-024-21

Query Match 59.3%; Score 2172.8; DB 3; Length 7333;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY	1479	GAGTTACCGCGGGCGGCTGGCGTAATCATGTGTATAGTCTTCTCTGTGAAATT	1538
DB	3576	GAGTCGACGGGATCAAGCTTGGCGTAATCATGTGTATAGTCTTCTCTGTGAAATT	3635
QY	1539	GTTATCCGCTCACAAATCCACACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGG	1598
DB	3636	GTTATCCGCTCACAAATCCACACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGG	3695
QY	1599	GTGCTAATAGTAGTAACTCAATTAAATGCGTTCGCTCACTGCGCGCTTTCCAGT	1658
DB	3696	GTGCTAATAGTAGTAACTCAATTAAATGCGTTCGCTCACTGCGCGCTTTCCAGT	3755
QY	1659	CGGGAACCTGTCGTCAGCTGCAATTAATGAATCGGCCACGCGCGGGAGAGCGGTT	1718
DB	3756	CGGGAACCTGTCGTCAGCTGCAATTAATGAATCGGCCACGCGCGGGAGAGCGGTT	3815
QY	1719	TGCGTATTGGGCGCTCTTCGCTTCCTCGCTCACTGACTCGCTGGGCTCGGTTCGGC	1778
DB	3816	TGCGTATTGGGCGCTCTTCGCTTCCTCGCTCACTGACTCGCTGGGCTCGGTTCGGC	3875
QY	1779	TGCGCGAGCGGTATCAGCTCACTCAAAAGCGGTAAATACGTTATCCACAGAAATCAGGG	1838
DB	3876	TGCGCGAGCGGTATCAGCTCACTCAAAAGCGGTAAATACGTTATCCACAGAAATCAGGG	3935
QY	1839	ATAACGAGGAAGAAACATGTGACAAAGGCCAGCAAAAGGCCAGAACCGTAAAGG	1898
DB	3936	ATAACGAGGAAGAAACATGTGACAAAGGCCAGCAAAAGGCCAGAACCGTAAAGG	3995
QY	1899	CGCGGTGCTGGCGTTTTCATAGGCTCCGCCCCCTTGACGAGCATCACAAAATCGAC	1958
DB	3996	CGCGGTGCTGGCGTTTTCATAGGCTCCGCCCCCTTGACGAGCATCACAAAATCGAC	4055
QY	1959	GCTCAAGTCAGAGTGGCGAAACCCGACAGACTATAAGATACAGCGGTTTCCCCCTG	2018
DB	4056	GCTCAAGTCAGAGTGGCGAAACCCGACAGACTATAAGATACAGCGGTTTCCCCCTG	4115
QY	2019	GAAGTCCCTCGTGGCTCTCTGTTCCGACCGCTGCGCTTACCGGATACCTGTCCGCT	2078
DB	4116	GAAGTCCCTCGTGGCTCTCTGTTCCGACCGCTGCGCTTACCGGATACCTGTCCGCT	4175

QY	2079	TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGG	2138
DB	4176	TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGG	4235
QY	2139	TGTAGGTGCTTCGCTCAAGCTGGCTGTGTGCAGAAACCCCGCTTACGCCGACCGCT	2198
DB	4236	TGTAGGTGCTTCGCTCAAGCTGGCTGTGTGCAGAAACCCCGCTTACGCCGACCGCT	4295
QY	2199	GGCCTTATCCGGTAACCTATCGTCTTGAGTCAACCCCGTAAAGACACGACTTATCGCCAC	2258
DB	4296	GGCCTTATCCGGTAACCTATCGTCTTGAGTCAACCCCGTAAAGACACGACTTATCGCCAC	4355
QY	2259	TGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT	2318
DB	4356	TGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT	4415
QY	2319	TCTTGAAGTGGTGGCTTAACCTACGGCTACACTAGAAAGACAGATATTTGGTATCTCGGCTC	2378
DB	4416	TCTTGAAGTGGTGGCTTAACCTACGGCTACACTAGAAAGACAGATATTTGGTATCTCGGCTC	4475
QY	2379	TGCTGAAGCCAGTTACCTTCGGAAGAAAGAGTTGGTAGCTCTTGATCCGCGAAACAAACCA	2438
DB	4476	TGCTGAAGCCAGTTACCTTCGGAAGAAAGAGTTGGTAGCTCTTGATCCGCGAAACAAACCA	4535
QY	2439	CCGCTGTAGCGGTGGTTTTTTTGTTCGCAAGCAGCAGATTACGCGCAGAAAGAAAGAT	2498
DB	4536	CCGCTGTAGCGGTGGTTTTTTTGTTCGCAAGCAGCAGATTACGCGCAGAAAGAAAGAT	4595
QY	2499	CTCAAGAAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGGAACGAAACTCAC	2558
DB	4596	CTCAAGAAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGGAACGAAACTCAC	4655
QY	2559	GTTAAGGATTTTGGTCAATGAGATTATCAAAAGAGTCTTCAACCTAGATCTCTTTTAAAT	2618
DB	4656	GTTAAGGATTTTGGTCAATGAGATTATCAAAAGAGTCTTCAACCTAGATCTCTTTTAAAT	4715
QY	2619	AAAAATGAAGTTTTAAATCAATTAAGTATATATAGTAAACTTTGGTCTGACAGTTACC	2678
DB	4716	AAAAATGAAGTTTTAAATCAATTAAGTATATATAGTAAACTTTGGTCTGACAGTTACC	4775
QY	2679	AATGCTTAATCAGTGGAGCCTATCTCAGCGATCTGTCTATTTGTTTCATCCATAGTTG	2738
DB	4776	AATGCTTAATCAGTGGAGCCTATCTCAGCGATCTGTCTATTTGTTTCATCCATAGTTG	4835
QY	2739	CCTGACTCCCGTCTGTGTAGATAACTACGATACGGGAGGCTTACCATCTGCGCCCAAGTG	2798
DB	4836	CCTGACTCCCGTCTGTGTAGATAACTACGATACGGGAGGCTTACCATCTGCGCCCAAGTG	4895
QY	2799	CTGCAATGATACCGGAGACCCAGCTCACCGGCTCCAGATTTTATCAGCAATAAAACAGC	2858
DB	4896	CTGCAATGATACCGGAGACCCAGCTCACCGGCTCCAGATTTTATCAGCAATAAAACAGC	4955
QY	2859	CAGCGGAAGGCGGAGCGGAGAGTGTCTGTGAACTTTTATCCGCTCCATCCAGTCTA	2918
DB	4956	CAGCGGAAGGCGGAGCGGAGAGTGTCTGTGAACTTTTATCCGCTCCATCCAGTCTA	5015
QY	2919	TTAATTGTTGCGGGAAGCTAGAGTAAGTAGTTCCGCCAGTTAAATAGTTTGGCAACGTTG	2978
DB	5016	TTAATTGTTGCGGGAAGCTAGAGTAAGTAGTTCCGCCAGTTAAATAGTTTGGCAACGTTG	5075
QY	2979	TTGCCATTGCTACAGGCATCGTGGTGTCACTGCTGCTTGGTATGCTTCATTCAGCT	3038
DB	5076	TTGCCATTGCTACAGGCATCGTGGTGTCACTGCTGCTTGGTATGCTTCATTCAGCT	5135
QY	3039	CGGTTTCCCAAGCATCAAGCGGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTA	3098
DB	5136	CGGTTTCCCAAGCATCAAGCGGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTA	5195
QY	3099	GCTCCTTCGGTCTCCGATCGTTGTGCAAGTAAGTTGGCGCGCAGTGTATCATTCACTGG	3158
DB	5196	GCTCCTTCGGTCTCCGATCGTTGTGCAAGTAAGTTGGCGCGCAGTGTATCATTCACTGG	5255
QY	3159	TTATGGCAGCAGTGCATAAATTTCTTCTTACTGTGATGCCCATCCGTAAGATGCTTTTCTGTGA	3218

Db	5256	TTATGCGAGACTGCATAATCTCTTACTGTGTCATGCCATCCGTAAGATGCTTTTCTGTGA	5315
Qy	3219	CTGGTGAGTACTCAACCAAGTCACTCTGAGAAATAGTGTATGCGGACCGAGTTGCTCTT	3278
Db	5316	CTGGTGAGTACTCAACCAAGTCACTCTGAGAAATAGTGTATGCGGACCGAGTTGCTCTT	5375
Qy	3279	GCCCGCGCTCAATACGGGATAATACCGCGCCACATAGCAAGACTTTAAAGTGTCTATCA	3338
Db	5376	GCCCGCGCTCAATACGGGATAATACCGCGCCACATAGCAAGACTTTAAAGTGTCTATCA	5435
Qy	3339	TTGGAAAACGTTCTTTCGGGGCGAAACTCTCAAGGATCTTACCGCTGTGTGAGATCCAGTT	3398
Db	5436	TTGGAAAACGTTCTTTCGGGGCGAAACTCTCAAGGATCTTACCGCTGTGTGAGATCCAGTT	5495
Qy	3399	CGATGTAAACCCACTCGTGCACCCAACTGCTTTCAGCATCTTTTACCTTTTACCAGCGTTT	3458
Db	5496	CGATGTAAACCCACTCGTGCACCCAACTGCTTTCAGCATCTTTTACCTTTTACCAGCGTTT	5555
Qy	3459	CTGGGTGAGCAAAAACAGGAAGCAAAATGCGCGCAAAAAGGGAATAAGGGCGACACGGA	3518
Db	5556	CTGGGTGAGCAAAAACAGGAAGCAAAATGCGCGCAAAAAGGGAATAAGGGCGACACGGA	5615
Qy	3519	AATGTTGAATACCTCATACTCTTCTTCTTTTCAATATTATTGAAGCATTTATCAGGGTTATT	3578
Db	5616	AATGTTGAATACCTCATACTCTTCTTCTTTTCAATATTATTGAAGCATTTATCAGGGTTATT	5675
Qy	3579	GTCTCATGAGCGGATACATATTGTAATGTAATTTAGAAAAATAAACAATAGGGTTCCGC	3638
Db	5676	GTCTCATGAGCGGATACATATTGTAATGTAATTTAGAAAAATAAACAATAGGGTTCCGC	5735
Qy	3639	GCACATTTCCCGAAAAGTGCCAC	3662
Db	5736	GCACATTTCCCGAAAAGTGCCAC	5759

RESULT 14

```

US-09-404-066-21
; Sequence 21, Application US/09404066
; Patent No. 6365409
; GENERAL INFORMATION:
; APPLICANT: Bradley, John D.
; APPLICANT: Thompson, Craig M.
; APPLICANT: Moore, Jeffrey B.
; APPLICANT: Wobbe, C. Richard
; APPLICANT: Healy, Judith M.
; APPLICANT: Donnelly, Caroline E.
; TITLE OF INVENTION: REGULATED GENE EXPRESSION IN YEAST
; FILE REFERENCE: 0342/10469051
; CURRENT APPLICATION NUMBER: US/09/404,066
; CURRENT FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/138,024
; PRIOR FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: 60/056,719
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 7333
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid p2M197
US-09-404-066-21

```

Query Match 59.3%; Score 2172.8; DB 3; Length 7333;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2177; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy	1479	GAGTTACCGCGGGCGCTTGGCGCTTAATCATGGTCATAGCTGTTTCTGTTGTAAT	1538
Db	3576	GAGTCACCGCGATGCAAGCTTGGCGTAAATCATGGTCATAGCTGTTTCTGTTGTAAT	3635

Qy	1539	GTATATCCGCTCACAATTCACACAACATACGAGCGGAGCATTAAGTGTAAAGCCTGGG	1598
Db	3636	GTATATCCGCTCACAATTCACACAACATACGAGCGGAGCATTAAGTGTAAAGCCTGGG	3695
Qy	1599	GTGCTTAATGAGTGAGCTAACTCACAATTAATGCGTTGCGCTCAGTCGCCGCTTTCCAGT	1658
Db	3696	GTGCTTAATGAGTGAGCTAACTCACAATTAATGCGTTGCGCTCAGTCGCCGCTTTCCAGT	3755
Qy	1659	CGGGAACCTGTCGTGCCAGCTGCATTAATGAATCGGCAACGCGCGGGAGAGCGGTT	1718
Db	3756	CGGGAACCTGTCGTGCCAGCTGCATTAATGAATCGGCAACGCGCGGGAGAGCGGTT	3815
Qy	1719	TGCGTATTGGGCGCTTTCGCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCCGC	1778
Db	3816	TGCGTATTGGGCGCTTTCGCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCCGC	3875
Qy	1779	TGCGGCGAGCGGTATCAGCTCACTCAAGGCGGTAAATACGGTTATCCACAGAAATCAGGGG	1838
Db	3876	TGCGGCGAGCGGTATCAGCTCACTCAAGGCGGTAAATACGGTTATCCACAGAAATCAGGGG	3935
Qy	1839	ATAACGCGAGGAAGAAACATGTGAGCAAAAGGCGCAAAAGGCGAGAAACCGTAAAAAGG	1898
Db	3936	ATAACGCGAGGAAGAAACATGTGAGCAAAAGGCGCAAAAGGCGAGAAACCGTAAAAAGG	3995
Qy	1899	CGCGCTTGTGCGGCTTTCATAGGCTCGCGCCCTCGACGAGCATCAAAAATCGAC	1958
Db	3996	CGCGCTTGTGCGGCTTTCATAGGCTCGCGCCCTCGACGAGCATCAAAAATCGAC	4055
Qy	1959	GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACAGGGGTTTCCCGCTG	2018
Db	4056	GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACAGGGGTTTCCCGCTG	4115
Qy	2019	GAAAGTCCCTCGTGCCTCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCT	2078
Db	4116	GAAAGTCCCTCGTGCCTCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCT	4175
Qy	2079	TTCTCCTTCGCGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCG	2138
Db	4176	TTCTCCTTCGCGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCG	4235
Qy	2139	TGTAGTGTCTCGCTCCAAAGCTGGGCTGTGCAGAACCCCGCTTACGCCGACCGCT	2198
Db	4236	TGTAGTGTCTCGCTCCAAAGCTGGGCTGTGCAGAACCCCGCTTACGCCGACCGCT	4295
Qy	2199	GCGCTTATCCGGTAACTATCGTCTTGTAGTCCAAACCCCGTAAAGACACGACTTATCCGAC	2258
Db	4296	GCGCTTATCCGGTAACTATCGTCTTGTAGTCCAAACCCCGTAAAGACACGACTTATCCGAC	4355
Qy	2259	TGCGAGCGCACTCGTAAACAGGATTTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAT	2318
Db	4356	TGCGAGCGCACTCGTAAACAGGATTTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAT	4415
Qy	2319	TCTTGAAGTGGTGGCTTAACCTACGCTACATAGAGGACAGTATTGTGTATCTGCGCTC	2378
Db	4416	TCTTGAAGTGGTGGCTTAACCTACGCTACATAGAGGACAGTATTGTGTATCTGCGCTC	4475
Qy	2379	TGCTGAAGCCAGTATACCTTCGGAAAAAGAGTTGGTAGCTCTTTCATCCGCCAAACCA	2438
Db	4476	TGCTGAAGCCAGTATACCTTCGGAAAAAGAGTTGGTAGCTCTTTCATCCGCCAAACCA	4535
Qy	2439	CGCGTGTAGCGGTGTTTTTTTGTTCGAAAGCAGATTTACCGCGAGAAAAAAGGAT	2498
Db	4536	CGCGTGTAGCGGTGTTTTTTTGTTCGAAAGCAGATTTACCGCGAGAAAAAAGGAT	4595
Qy	2499	CTCAAGAGATCCTTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTC	2558
Db	4596	CTCAAGAGATCCTTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTC	4655
Qy	2559	GTTAAGGGATTTTGTGTGATGATTAATCAAAAGGATCTTACCTAGATCTTTTAAAT	2618
Db	4656	GTTAAGGGATTTTGTGTGATGATTAATCAAAAGGATCTTACCTAGATCTTTTAAAT	4715
Qy	2619	AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC	2678

|||||
4716 AAAAAAGAAAGTTTAAATCAATCTAAAGTATATATAGTAAAGCTTGGTCTGACAGTTACC 4775
QY AATGCTTAATCAGTGAAGGACCTATCTCAGCGATCTGTCTATTTCGTTCCATCCATAGTTG 2738
Db AATGCTTAATCAGTGAAGGACCTATCTCAGCGATCTGTCTATTTCGTTCCATCCATAGTTG 4835
QY CTTGACTCCCCGTCGTGTAGATAAATACATACGGGAGGGCTTACCATCTGGCCCCCAGTG 2798
Db CTTGACTCCCCGTCGTGTAGATAAATACATACGGGAGGGCTTACCATCTGGCCCCCAGTG 4895
QY CTGCAATGATACCGGAGACCCACGCTCACCGGCTCCAGATTATCAGCAATAAACCCAGC 2858
Db CTGCAATGATACCGGAGACCCACGCTCACCGGCTCCAGATTATCAGCAATAAACCCAGC 4955
QY CAGCCGGAAGGCCGAGCGCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTA 2918
Db CAGCCGGAAGGCCGAGCGCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTA 5015
QY TTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTAAATAGTTTGGCCAAAGTTG 2978
Db TTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTAAATAGTTTGGCCAAAGTTG 5075
QY TTGCCATTGCTACAGCATCTGTTGGTGTACGCTCGTGGTTCGTTGCTATGCTTCATTCACT 3038
Db TTGCCATTGCTACAGCATCTGTTGGTGTACGCTCGTGGTTCGTTGCTATGCTTCATTCACT 5135
QY CCGGTTCCCAACGATCAAGCGAGTACATGATCCCCATGTTGTGCAAAAAGCGGTTA 3098
Db CCGGTTCCCAACGATCAAGCGAGTACATGATCCCCATGTTGTGCAAAAAGCGGTTA 5195
QY GCTCCTTCGGTCTCCGATCGTTGTGAGAAAGTATGAGTGGCCGAGTGTATCACTCATGG 3158
Db GCTCCTTCGGTCTCCGATCGTTGTGAGAAAGTATGAGTGGCCGAGTGTATCACTCATGG 5255
QY TTATGGCAGCATGCATAATCTCTTACTGTGATGTCATGCGCATTCGTAAGATGCTTTCTGTA 3218
Db TTATGGCAGCATGCATAATCTCTTACTGTGATGTCATGCGCATTCGTAAGATGCTTTCTGTA 5315
QY CTGGTGAGTACTCAACCAAGTCAATCTGAGATAAGTATGATGCGGACCGAGTGTCTCTT 3278
Db CTGGTGAGTACTCAACCAAGTCAATCTGAGATAAGTATGATGCGGACCGAGTGTCTCTT 5375
QY GCCCGCGCTCAATACGGGATAATACCGGCCACATAGCAGAACTTTAAAGTGTCTATCA 3338
Db GCCCGCGCTCAATACGGGATAATACCGGCCACATAGCAGAACTTTAAAGTGTCTATCA 5435
QY TTGGAAGACGTTCTTTCGGGGCGAAATCTCTCAAGATCTTACCGTGTGTGAGATCCAGTT 3398
Db TTGGAAGACGTTCTTTCGGGGCGAAATCTCTCAAGATCTTACCGTGTGTGAGATCCAGTT 5495
QY CGATGTAAACCACTCGTGACCCCACTGATCTTTCAGCATCTTTTACTTTTCAACGCGTTT 3458
Db CGATGTAAACCACTCGTGACCCCACTGATCTTTCAGCATCTTTTACTTTTCAACGCGTTT 5555
QY CTGGGTGAGCAAAAACAGGAAGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGA 3518
Db CTGGGTGAGCAAAAACAGGAAGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGA 5615
QY AATGTTGAATACATCATCTCTTCTTTTCAATATATTGAGAGCAATTTATCAGGGTTATT 3578
Db AATGTTGAATACATCATCTCTTCTTTTCAATATATTGAGAGCAATTTATCAGGGTTATT 5675
QY GTCTCATGCGGATACATATTTGATGTAATTTAGAAAATAAACAATAGGGTTCCGC 3638
Db GTCTCATGCGGATACATATTTGATGTAATTTAGAAAATAAACAATAGGGTTCCGC 5735
QY GCACATTTCCCGAAAAGTGCCAC 3662
Db GCACATTTCCCGAAAAGTGCCAC 5759

RESULT 15

US-09-573-322-21
; Sequence 21, Application US/09573322
; Patent No. 6531289
; GENERAL INFORMATION:
; APPLICANT: Bradley, John D.
; APPLICANT: Thompson, Craig M.
; APPLICANT: Moore, Jeffrey B.
; APPLICANT: Wobbe, C. Richard
; APPLICANT: Bailey, David A.
; TITLE OF INVENTION: Regulated Gene Expression in Yeast and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 0342/ID469-US4
; CURRENT APPLICATION NUMBER: US/09/573,322
; CURRENT FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/404,066
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/138,024
; PRIOR FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: 60/056,719
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 7333
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ZM197 plasmid
US-09-573-322-21

Query Match 59.3%; Score 2172.8; DB 4; Length 7333;
Best Local Similarity 99.7%; Pred. No. 0;

Matches 2177; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1479 GAGTTACCGCGCGCGCTTGGCGTAATCATGTCTAGTGTCTTCTGTGTGAATTT 1538
Db 3576 GAGTCAGCGCGATCAAGCTTGGCGTAATCATGTCTAGTGTCTTCTGTGTGAATTT 3635
QY 1539 GTTATCCGCTCACATTCACACATACGAGCGGAGCATAAAGTGTAAAGCTGGG 1598
Db 3636 GTTATCCGCTCACATTCACACATACGAGCGGAGCATAAAGTGTAAAGCTGGG 3695
QY 1599 GTGCCATAATGAGTGAAGTCACTTAATTTGCGTTTGGCTCACTGCGCGTTTCCAGT 1658
Db 3696 GTGCCATAATGAGTGAAGTCACTTAATTTGCGTTTGGCTCACTGCGCGTTTCCAGT 3755
QY 1659 CGGAAACCTGTCGTCAGCTGCATTAATGAATCGGCAACCGCGGAGAGCGGTT 1718
Db 3756 CGGAAACCTGTCGTCAGCTGCATTAATGAATCGGCAACCGCGGAGAGCGGTT 3815
QY 1719 TGCATTATTTGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTCGCTCGC 1778
Db 3816 TGCATTATTTGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTCGCTCGC 3875
QY 1779 TCGCGAGCGGTTATCAGTCACTCAAAAGCGGTAATACGGTTATPCCAAGAAATCAGGG 1838
Db 3876 TCGCGAGCGGTTATCAGTCACTCAAAAGCGGTAATACGGTTATPCCAAGAAATCAGGG 3935
QY 1839 ATAACGAGGAAAGAAATGATGAGCAAAAGCGCAGCAAAAGCGCAGGAAACCGTAAAGG 1898
Db 3936 ATAACGAGGAAAGAAATGATGAGCAAAAGCGCAGCAAAAGCGCAGGAAACCGTAAAGG 3995
QY 1899 CCGGTTCTCGCGCTTTTTCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCCAG 1958
Db 3996 CCGGTTCTCGCGCTTTTTCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCCAG 4055
QY 1959 GCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACAGCGGTTTCCCCCTG 2018
Db 4056 GCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACAGCGGTTTCCCCCTG 4115
QY 2019 GAAGCTCCCTCGCGCTCTCTCTGTCGACCCCTGCGGCTTACCGGATACCTGTCCGCT 2078
Db 4116 GAAGCTCCCTCGCGCTCTCTCTGTCGACCCCTGCGGCTTACCGGATACCTGTCCGCT 4175

QY	2079	TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCG	2138
Db	4176	TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCG	4235
QY	2139	TGTAGGTCTGTCTCCTCAAGCTGGGCTGTGTGACGAACCCCGCTTCAGCCGACCGCT	2198
Db	4236	TGTAGGTCTGTCTCCTCAAGCTGGGCTGTGTGACGAACCCCGCTTCAGCCGACCGCT	4295
QY	2199	GGCCCTTATCCGGTAACTATCGTTTGGAGTCCAAACCCGGTAAAGACACGACTTATCGCCAC	2258
Db	4296	GGCCCTTATCCGGTAACTATCGTTTGGAGTCCAAACCCGGTAAAGACACGACTTATCGCCAC	4355
QY	2259	TGGCAGCAGCCACTGTTACAGGATTTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAT	2318
Db	4356	TGGCAGCAGCCACTGTTAAACAGGATTTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAT	4415
QY	2319	TCTTGAAGTGGTGGCCCTAACTACGGCTACACTAGAGGACAGTATTTGGTGTCTGGGCTC	2378
Db	4416	TCTTGAAGTGGTGGCCCTAACTACGGCTACACTAGAGGACAGTATTTGGTGTCTGGGCTC	4475
QY	2379	TGCTGAAGCCAGTTACTTTTCGGAAGAGTGGTAGCTCTTGATCCGGCAACCAACCA	2438
Db	4476	TGCTGAAGCCAGTTACTTTTCGGAAGAGTGGTAGCTCTTGATCCGGCAACCAACCA	4535
QY	2439	CGCTGGTAGCGGTGGTTTTTTTGGTGAAGCAGAGATTACGGCGCAGAAAGGAT	2498
Db	4536	CGCTGGTAGCGGTGGTTTTTTTGGTGAAGCAGAGATTACGGCGCAGAAAGGAT	4595
QY	2499	CTCAAGAGATCCTTTGATCTTTCTACGGGCTCTGACGCTCAGTGGAGCGAAACTCAC	2558
Db	4596	CTCAAGAGATCCTTTGATCTTTCTACGGGCTCTGACGCTCAGTGGAGCGAAACTCAC	4655
QY	2559	GTTAAGGATTTTGGTTCATGAGATTATCAAAAGGATCTTCACTAGATCTTTTAAATT	2618
Db	4656	GTTAAGGATTTTGGTTCATGAGATTATCAAAAGGATCTTCACTAGATCTTTTAAATT	4715
QY	2619	AAAAATGAAGTTTAAATCAATCTAAAGTATATAGTAAACTTGGTCTGACAGTTACC	2678
Db	4716	AAAAATGAAGTTTAAATCAATCTAAAGTATATAGTAAACTTGGTCTGACAGTTACC	4775
QY	2679	AATGCTTAATCAGTGAGGACCTATCTCAGCGATCTGTCTATTTCGTTCCATCCATAGTTG	2738
Db	4776	AATGCTTAATCAGTGAGGACCTATCTCAGCGATCTGTCTATTTCGTTCCATCCATAGTTG	4835
QY	2739	CCTGACTCCCGTCTGTAGATAACTACGATACGGGAGGCTTACCATCTGGCCCCAGTG	2798
Db	4836	CCTGACTCCCGTCTGTAGATAACTACGATACGGGAGGCTTACCATCTGGCCCCAGTG	4895
QY	2799	CTGCAATGATACCGGAGACCCACGCTCACCGGCTCCAGATTATCAGCAATAAACCCAGC	2858
Db	4896	CTGCAATGATACCGGAGACCCACGCTCACCGGCTCCAGATTATCAGCAATAAACCCAGC	4955
QY	2859	CAGCCGGAAGCGCGAGCGCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTA	2918
Db	4956	CAGCCGGAAGCGCGAGCGCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTA	5015
QY	2919	TTAATTTGTTGGCGGAAGCTAGAGTAAGTAGTTTCGCCAGTAAATAGTTTGGCAAGCTTG	2978
Db	5016	TTAATTTGTTGGCGGAAGCTAGAGTAAGTAGTTTCGCCAGTAAATAGTTTGGCAAGCTTG	5075
QY	2979	TTGCCATTGCTACAGGCATCGTGGTGTCA CGCTCGTGGTTGGTTCATTTCAGCT	3038
Db	5076	TTGCCATTGCTACAGGCATCGTGGTGTCA CGCTCGTGGTTGGTTCATTTCAGCT	5135
QY	3039	CGGTTCCCAACGATCAAGCGAGTTACATGATCCCGCTAGTTGTGCAAAAAAGCGGTTA	3098
Db	5136	CGGTTCCCAACGATCAAGCGAGTTACATGATCCCGCTAGTTGTGCAAAAAAGCGGTTA	5195
QY	3099	GCTCTTCGGTCTCCGATCGTTGTGAGAAAGTGGCGCAGTGGTTATCACTCATGG	3158
Db	5196	GCTCTTCGGTCTCCGATCGTTGTGAGAAAGTGGCGCAGTGGTTATCACTCATGG	5255

QY	3159	TTATGGCAGCACTGCATTAATTTCTTACTGTCTATGCCATCCGTAAAGATGCTTTTCTGTGA	3218
Db	5256	TTATGGCAGCACTGCATTAATTTCTTACTGTCTATGCCATCCGTAAAGATGCTTTTCTGTGA	5315
QY	3219	CTGGTGAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGCGACCGAGTTGCTCTT	3278
Db	5316	CTGGTGAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGCGACCGAGTTGCTCTT	5375
QY	3279	GCCCGCGTCAATACGGGATAATAACCGGCCACATAGCAGAACTTTTAAAGTCTCATCA	3338
Db	5376	GCCCGCGTCAATACGGGATAATAACCGGCCACATAGCAGAACTTTTAAAGTCTCATCA	5435
QY	3339	TTGGAAAAAGTCTTTCGGGGCGAAGAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT	3398
Db	5436	TTGGAAAAAGTCTTTCGGGGCGAAGAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT	5495
QY	3399	CGATGTAAACCCACTCGTGACCACTGATCTTTCAGCATCTTTTACTTTTACCAGCGTTT	3458
Db	5496	CGATGTAAACCCACTCGTGACCACTGATCTTTCAGCATCTTTTACTTTTACCAGCGTTT	5555
QY	3459	CTGGTGAGCAAAAAACAGGAAGCAAAATGCGCGCAAAAAAGGAAATAAGGGCGACACGGA	3518
Db	5556	CTGGTGAGCAAAAAACAGGAAGCAAAATGCGCGCAAAAAAGGAAATAAGGGCGACACGGA	5615
QY	3519	AATGTTGAATACTCATACTCTTCTTCTTTTCAATATTTTGAAGCATTTATCAGGGTTATT	3578
Db	5616	AATGTTGAATACTCATACTCTTCTTCTTTTCAATATTTTGAAGCATTTATCAGGGTTATT	5675
QY	3579	GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAGGGGTTCCGC	3638
Db	5676	GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAGGGGTTCCGC	5735
QY	3639	GCACATTTTCCCGAAAAAGTGCCAC	3662
Db	5736	GCACATTTTCCCGAAAAAGTGCCAC	5759

Search completed: June 5, 2005, 23:08:18
Job time : 656.194 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 13:38:04 ; Search time 10710.2 Seconds
(without alignments)
13014.775 Million cell updates/sec

Title: US-10-826-523-41

Perfect score: 3662

Sequence: 1 ctaaatgtaagcgtaata.....attccccgaaagtgcac 3662

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_hic.*
- 4: gb_est3.*
- 5: gb_est4.*
- 6: gb_est5.*
- 7: gb_est6.*
- 8: gb_gss1.*
- 9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1055.8	28.8	1070	1	AJ281552
2	1018.2	27.8	1048	7	CO552396
3	986.4	26.9	1013	4	BM438846
4	955.2	26.1	1067	1	AU081137
5	924	25.2	1004	1	AJ281480
6	923.8	25.2	1089	1	AU081124
7	902.8	24.7	1049	9	CL021189
8	889.8	24.3	928	7	CO487414
9	881.4	24.1	925	6	CB686151
10	876.4	23.9	902	7	CR753463
11	870.2	23.8	935	4	BG382279
12	865	23.6	1073	7	CF269652
13	857.6	23.4	885	9	CL076016
14	841	23.0	841	1	AL042026
15	824.6	22.5	854	4	BM438950
16	824	22.5	1169	9	AG332951
17	821.4	22.4	1025	9	CL021193
18	820.2	22.4	865	7	CK125894
19	819.6	22.4	1249	8	BZ572284
20	814.4	22.2	1126	8	BZ577702
21	809.4	22.1	856	7	CN823189
22	808	22.1	819	6	CD649375
23	807.6	22.1	966	8	BZ570738
24	797.6	21.8	1011	8	BZ576726

C 25	793	21.7	827	7	CN823902	Oa splbn
C 26	791	21.6	820	7	CR753457	DKFZp469P
C 27	788.6	21.5	833	6	CB686421	Bn01b_04j
C 28	786.4	21.5	1163	1	AU081044	AU081044
C 29	785.4	21.4	1574	8	BZ572566	msb2_2693
C 30	779.8	21.3	846	7	CV468077	est_I_van
C 31	777.8	21.2	800	1	AJ281449	4A3A-P4D5
C 32	774.8	21.2	1336	8	BZ575810	msb2_4637
C 33	774.4	21.1	789	6	CD280920	G44224.42
C 34	772.4	21.1	954	1	AL044364	DKFZp434C
C 35	771.4	21.1	832	7	CN822433	Oa splbn
C 36	767.6	21.0	863	7	CF752100	TGDR9_Hum
C 37	767.6	21.0	1370	8	BZ571721	msb2_2025
C 38	765.8	20.9	769	7	CV224987	CS_hyp_24
C 39	763.2	20.8	780	5	BQ825693	1030129B0
C 40	757	20.7	759	6	CD279661	G43818.35
C 41	755.4	20.6	1463	8	BZ571475	msb2_1906
C 42	752.4	20.5	1369	8	BZ579291	msb2_6285
C 43	751.4	20.5	786	7	CN823164	Oa splbn
C 44	750.4	20.5	752	6	CD646006	SSF526R
C 45	749	20.5	752	7	CR766850	DKFZp468H

ALIGNMENTS

RESULT 1	AJ281552	1070 bp	mRNA	linear	EST 30-JUN-2000
LOCUS	4A3A-P6F11-F	Anopheles gambiae	immune competent	4A3A	Anopheles
DEFINITION	gambiae cDNA clone 4A3A-P6F11, mRNA sequence.				
ACCESSION	AJ281552				
VERSION	AJ281552.1	GI:6929432			
KEYWORDS	EST.				
SOURCE	Anopheles gambiae	(African malaria mosquito)			
ORGANISM	Anopheles gambiae				
REFERENCE	1 (bases 1 to 1070)				
AUTHORS	Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M.B. and Kafatos, F.C.				
TITLE	Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	97 (12)	6619-6624	(2000)	
MEDLINE	20300950				
PUBMED	10841561				
COMMENT	Contact: Dimopoulos G Fotis C. Kafatos laboratory European Molecular Biology Laboratory Meyerhofstrasse 1, 69117 Heidelberg, Germany.				
FEATURES	Location/Qualifiers				
source	i. .1070				
	/organism="Anopheles gambiae"				
	/mol_type="mRNA"				
	/strain="4A r/r"				
	/db_xref="taxon:7165"				
	/clone="4A3A-P6F11"				
	/cell_line="immune competent 4A3A"				
	/lab_host="E. coli DH10B"				
	/clone_lib="Anopheles gambiae immune competent 4A3A"				
	/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: EcoRI; Site 2: NotI; sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996) : Normalization and Subtraction: Two approaches To Facilitate Gene Discovery, Genome Research 6, 791-806."				
ORIGIN					

Query Match		28.8%;	Score 1055.8;	DB 1;	Length 1070;
Best Local Similarity		99.7%;	Pred. No. 3.9e-299;		
Matches 1068;		Conservative	0;	Mismatches	2; Indels 1; Gaps 1;
QY	1904	TTGCTGGCGT	TTTTCCATAGGCTCGCC	CCCCCTGACGAGCATC	CAAAAAATCGAGCTCA 1963
DB	1	TTGCTGGCGT	TTTTCCATAGGCTCGCC	CCCCCTGACGAGCATC	CAAAAAATCGAGCTCA 60
QY	1964	AGTCAGAGTGGCGAA	CCCGACAGACTAT	ATAAGATAC	CAGCGCTTTCCCTCGGAAGC 2023
DB	61	AGTCAGAGTGGCGAA	CCCGACAGACTAT	ATAAGATAC	CAGCGCTTTCCCTCGGAAGC 120
QY	2024	TCCTCGTGGCTCT	CTCTGTTCCGACCCCTGCGCTT	ACCGGATAC	CTGTCGCCCTTTCTC 2083
DB	121	TCCTCGTGGCTCT	CTCTGTTCCGACCCCTGCGCTT	ACCGGATAC	CTGTCGCCCTTTCTC 180
QY	2084	CTTCCGGAAGCGTGG	CGCTTTCTCATAGCTCAC	CGCTGTAGTATCT	CAGTTCGGTGTAG 2143
DB	181	CTTCCGGAAGCGTGG	CGCTTTCTCATAGCTCAC	CGCTGTAGTATCT	CAGTTCGGTGTAG 240
QY	2144	GTTCGTTCCGCTCA	AGCTGGGCTGTGTGC	CAGAACCCCGCTT	CAGCCGACCGCTGCGCC 2203
DB	241	GTTCGTTCCGCTCA	AGCTGGGCTGTGTGC	CAGAACCCCGCTT	CAGCCGACCGCTGCGCC 300
QY	2204	TTATCCGGTAACTA	TCGTCTTGAGTCCAA	CCCGGTAAGACAC	AGCACTTATCGCACTGGCA 2263
DB	301	TTATCCGGTAACTA	TCGTCTTGAGTCCAA	CCCGGTAAGACAC	AGCACTTATCGCACTGGCA 360
QY	2264	CGAGCACTGGTAA	CAGATTTAGCAGAG	GAGGTATGTAG	GGCGTGTCTACAGATTCCTG 2323
DB	361	CGAGCACTGGTAA	CAGATTTAGCAGAG	GAGGTATGTAG	GGCGTGTCTACAGATTCCTG 420
QY	2324	AAGTGGTGGCTTA	CTACGGCTACACTA	GAGACAGATTT	TGTTGTTATCTGGCTCTGCTG 2383
DB	421	AAGTGGTGGCTTA	CTACGGCTACACTA	GAGACAGATTT	TGTTGTTATCTGGCTCTGCTG 480
QY	2384	AAGCCAGTTACCT	TCGGAAGAGT	TGGTAGCTCT	TGATCGGCGTGTACAGATTCCTG 2443
DB	481	AAGCCAGTTACCT	TCGGAAGAGT	TGGTAGCTCT	TGATCGGCGTGTACAGATTCCTG 540
QY	2444	GGTAGCGGTGGT	TTTTTTTGTGCAAG	CAGCAGATTT	ACGGCAGAAAAAAGGATCTCAA 2503
DB	541	GGTAGCGGTGGT	TTTTTTTGTGCAAG	CAGCAGATTT	ACGGCAGAAAAAAGGATCTCAA 600
QY	2504	GAAAGATCTTTG	ATCTTTTACGGGCT	CTGACGCTC	AGTGGAAACCAAACTCAGCTTAA 2563
DB	601	GAAAGATCTTTG	ATCTTTTACGGGCT	CTGACGCTC	AGTGGAAACCAAACTCAGCTTAA 660
QY	2564	GGGATTTTGGTCA	TAGATTTATCAAAA	AGGATCTTT	CACCTAGATCCTTTTAAATTTAAAA 2623
DB	661	GGGATTTTGGTCA	TAGATTTATCAAAA	AGGATCTTT	CACCTAGATCCTTTTAAATTTAAAA 720
QY	2624	TGAAGTTTAAAT	CTAATCTAAGTAT	TATATAGTAA	AACTTGGTCTGACAGTTTACCAATGC 2683
DB	721	TGAAGTTTAAAT	CTAATCTAAGTAT	TATATAGTAA	AACTTGGTCTGACAGTTTACCAATGC 780
QY	2684	TTAATCAGTAGG	CACTTCTCAGGAT	CTGTCTATTT	CGTTCCATCTCCTAGTGGCTCTGA 2743
DB	781	TTAATCAGTAGG	CACTTCTCAGGAT	CTGTCTATTT	CGTTCCATCTCCTAGTGGCTCTGA 840
QY	2744	CTCCCGCTCGT	GTAGATAACTAC	GATACGGAGG	GGCTTTACCATCTGGCCCCCAGTGTGCA 2803
DB	841	CTCCCGCTCGT	GTAGATAACTAC	GATACGGAGG	GGCTTTACCATCTGGCCCCCAGTGTGCA 900
QY	2804	ATGATACCGG	CAGACCCAGCT	CACCGCTC	CAGATTTTATCAGCAATTAACACGACGCC 2863
DB	901	ATGATACCGG	CAGACCCAGCT	CACCGCTC	CAGATTTTATCAGCAATTAACACGACGCC 960
QY	2864	GGAAGGCGG	CAGCAGAGT	GGTCTCTCA	CTTTATCCGCTCCATCCAGTCTATTAT 2923
DB	961	GGAA-GGCGG	CAGCAGAGT	GGTCTCTCA	CTTTATCCGCTCCATCCAGTCTATTAT 1019
QY	2924	TGTTGCCGG	GAAGCTAG	AGTATTG	TCGCCAGTTAATAGTTTTCGCCAAC 2974
DB					

1020

TGTTCCCGGAGCTAGAGTAAGTAGTTCCGCAAGTTAATAGTTTCGCCAAC

1070

RESULT 2

COS52396

LOCUS

DEFINITION

ACly4.50 Sea lamprey Acly Petromyzon marinus cDNA, mRNA sequence.

ACCESSION

COS52396

VERSION

COS52396.1

GI:51800732

KEYWORDS

EST.

SOURCE

Petromyzon marinus (sea lamprey)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia; Petromyzontiformes; Petromyzontidae; Petromyzon.

REFERENCE

1 (bases 1 to 1048)

AUTHORS

Pancer,Z., Mayer,W.E., Klein,J. and Cooper,M.D.

TITLE

Prototypic T-cell receptor and CD4-like coreceptor expressed in lymphocytes of the agnathan sea lamprey

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 101, 13273-13278 (2004)

COMMENT

Contact: Pancer, Zeev

Division of Developmental and Clinical Immunology

The University of Alabama at Birmingham

378 Wallace Tumor Institute, 1530 Third Avenue, South, Birmingham, AL 35294-3300

Tel: 205-975-5812

Fax: 205-975-7218

Email: zpancer@uab.edu

location/Qualifiers

1..1048

/organism="Petromyzon marinus"

/mol_type="mRNA"

/db_xref="taxon:7757"

/cell_type="lymphocyte"

/dev_stage="immune stimulated larvae"

/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR subtracted cDNA libraries of immune stimulated larvae. All are single pass 5' or 3' sequences randomly cloned in pGEM-T Easy (Promega)."

FEATURES

source

Query Match

Best Local Similarity

27.8%;

Score 1018.2;

DB 7;

Length 1048;

Matches 1043;

Conservative

1;

Mismatches

4;

Indels

3;

Gaps

2;

QY

1840

TAA

CGCAGAAAGACATGT

GCAGCAAAAGGCGCAGCAAAAGGCGCAGGACCGTTAAAGGCG

1899

DB

1

TAA

CGAAGGAAGAACATGT

GCAGCAAAAGGCGCAGCAAAAGGCGCAGGACCGTTAAAGGCG

60

QY

1900

CGCGTTGCTGGCGT

TTTTTCCATAGGCTCCGCCCCCTCAGCAGCATCA

CAAAAAATCGACG

1959

DB

61

CGCGTTGCTGGCGT

TTTTTCCATAGGCTCCG-CCCCCTGACGAGCATCA

CAAAAAATCGAGC

119

QY

1960

CTCAAGTCAGAGTGG

CGAAACCCGACAGGACTATAAGATACCAAGCGCTTTCC

CCCCCTGG

2019

DB

120

CTCAAGTCAGAGTGG

CGAAACCCGACAGGACTATAAGATACCAAGCGCTTTCC

CCCCCTGG

179

QY

2020

AAGCTCCCTCGTGG

CTCTCTGTTCCGACCTCGCGCTTACCGATACCTGTCGCGCTT

2079

DB

180

AAGCTCCCTCGTGG

CTCTCTGTTCCGACCTCGCGCTTACCGATACCTGTCGCGCTT

239

QY

2080

TCTCCCTTCGGAA

CGAGCGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGT

2139

DB

240

TCTCCCTTCGGAA

AGSGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGT

299

QY

2140

GTAAGTCGTT

CGCTTCAAGCTGGGCTGTGTGACGAAACCCCGCTTCAGCCGACCGGTG

2199

DB

300

GTAAGTCGTT

CGCTTCAAGCTGGGCTGTGTGACGAAACCCCGCTTCAGCCGACCGGTG

359

QY

2200

CGCCTTATCCG

TAACCTATCGTTGAGTCCAAACCCGCTTAAGACACGACTTATCGCCACT

2259

DB

360

CGCCTTATCCG

TAACCTATCGTTGAGTCCAAACCCGCTTAAGACACGACTTATCGCCACT

419

```
QY 2260 GGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGTATGTAAGCGGTGCTACAGAGTT 2319
Db 420 GGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGTATGTAAGCGGTGCTACAGAGTT 479
QY 2320 CTTTGAAGTGGTGGCCCTAACTACCGCTACACTAGAAGCAGATATTTGGTATCTGCGCTCT 2379
Db 480 CTTTGAAGTGGTGGCCCTAACTACCGCTACACTAGAAGCAGATATTTGGTATCTGCGCTCT 539
QY 2380 GCTGAAGCCAGTTACCTTCGAAAAAGAGTTGTGTAGCTCTTGATCCGGCAAAACAAACCCAC 2439
Db 540 GCTGAAGCCAGTTACCTTCGAAAAAGAGTTGTGTAGCTCTTGATCCGGCAAAACAAACCCAC 599
QY 2440 CGCTGTAGCGGTGTTTGTGTTGCAAGCAGCAGATTAACCGCAGAAAAAAGGATC 2499
Db 600 CGCTGTAGCG--TGTTTTTTTGTGCAAGCAGCAGATTAACCGCAGAAAAAAGGATC 657
QY 2500 TCAAGAGAGATCCTTTTCATCTTTCTACGGGTCTGACGCTCAGTGGAAACGAAACCTCAGG 2559
Db 658 TCAAGAGAGATCCTTTTCATCTTTCTACGGGTCTGACGCTCAGTGGAAACGAAACCTCAGG 717
QY 2560 TTAAGGGATTTTGGTTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTA 2619
Db 718 TTAAGGGATTTTGGTTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTA 777
QY 2620 AAAATGAAGTTTAAATCAATTAAGTATATATAGTAAACTTTGGTCTGACAGTTACCA 2679
Db 778 AAAATGAAGTTTAAATCAATTAAGTATATATAGTAAACTTTGGTCTGACAGTTACCA 837
QY 2680 ATGCTTAATCAGTGGAGCCATCTCAGGATCTCTATTTCTGTTTCCATCATAGTTGC 2739
Db 838 ATGCTTAATCAGTGGAGCCATCTCAGGATCTCTATTTCTGTTTCCATCATAGTTGC 897
QY 2740 CTGACTCCCGCTCGTGTAGATACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC 2799
Db 898 CTGACTCCCGCTCGTGTAGATACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC 957
QY 2800 TGAATGATACCGCGAGACCCAGCTCAGCGTCCAGATTTATACGCAATTAACCAAGCC 2859
Db 958 TGAATGATACCGCGAGACCCAGCTCAGCGTCCAGATTTATACGCAATTAACCAAGCC 1017
QY 2860 AGCCGAAGGGCGGAGCGCAGAGTGTCT 2890
Db 1018 AGCCGAAGGGCGGAGCGCAGAGTGTCT 1048
```

RESULT 3

```
BM438846/c
LOCUS 1pLvr00157 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA
DEFINITION sequence.
ACCESSION BM438846
VERSION BM438846.1 GI:18460568
KEYWORDS EST.
SOURCE Ictalurus punctatus (channel catfish)
ORGANISM Ictalurus punctatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
Ictaluridae; Ictalurus.
1 (bases 1 to 1013)
Feng, J., Kucuktas, H., Kocabas, A., Li, P. and Liu, Z.
Transcriptome of channel catfish (Ictalurus punctatus): initial
analysis of expressed sequence tags from the liver
Unpublished (2002)
Contact: Liu ZJ
The Fish Molecular Genetics and Biotechnology Laboratory,
Department of Fisheries and Allied Aquacultures and Program of Cell
and Molecular Biosciences
Auburn University
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
Tel: 334 844 4054
Fax: 334 844 9208
Email: zliu@acesag.auburn.edu
```

```
Seq primer: M13 Reverse.
Location/Qualifiers
1. .1013
/organism="Ictalurus punctatus"
/mol_type="mRNA"
/db_xref="taxon:7998"
/clone_lib="Liver cDNA library"
/note="Organ: Liver; Vector: pSport1; Site_1: NotI;
Site_2: Sall"

ORIGIN
Query Match 26.9%; Score 986.4; DB 4; Length 1013;
Best Local Similarity 99.5%; Pred. No. 1.1e-278;
Matches 1009; Conservative 1; Mismatches 2; Indels 2; Gaps 2;
QY 2477 ATTACCGCGCAGAAAAAAGGATCTCAAGAAAGATCTCTTGGATCTTTTCTACGGGTCTGAC 2536
Db 1013 ATTACCGCGCAGAAAAAAGGATCTCAAGAAAGATCTCTTGGATCTTTTCTACGGGTCTGAC 954
QY 2537 GCTCAGTGAACGAAACCTCAGCTTAAGGGATTTTGGTCATGAGATATCAAAAAGATC 2596
Db 953 GCTCAGTGAACGAAACCTCAGCTTAAGGGATTTTGGTCATGAGATATCAAAAAGATC 894
QY 2597 TTCACCTAGATCCTTTTAAATTAAGGATTTTAAATCAATCTAAAGTATATATGAG 2656
Db 893 TTCACCTAGATCCTTTTAAATTAAGGATTTTAAATCAATCTAAAGTATATATGAG 834
QY 2657 TAAACTTGTCTGACAGTTACCAATCTTAAATCAGTGAGGCACCTATCTCAGCGATCTGT 2716
Db 833 TAAACTTGTCTGACAGTTACCAATCTTAAATCAGTGAGGCACCTATCTCAGCGATCTGT 774
QY 2717 CTATTTGCTTCATCCATAGTTGCTGATCCCGCTCGTGTAGATAACTACGATACGGGAG 2776
Db 773 CTATTTGCTTCATCCATAGTTGCTGATCCCGCTCGTGTAGATAACTACGATACGGGAG 714
QY 2777 GGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGAGACCCACGCTCACCGGTCCA 2836
Db 713 GGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGAGACCCACGCTCACCGGTCCA 654
QY 2837 GATTTATCAGCAATAAACCCAGCCAGCGGCGGAGAGAGTGTCTCTGCAACT 2896
Db 653 GATTTATCAGCAATAAACCCAGCCAGCGGCGGAGAGAGTGTCTCTGCAACT 594
QY 2897 TTATCCGCTCCATCCAGTCTATTAATTTGTGCGGGAAGCTAGAGTAAGTAGTTCGCCA 2956
Db 593 TTATCCGCTCCATCCAGTCTATTAATTTGTGCGGGAAGCTAGAGTAAGTAGTTCGCCA 534
QY 2957 GTTAATAGTTTGGCAACGTTGTGTCATTTGCTACAGGCATCGTGTGTCAAGCTCGTCG 3016
Db 533 GTTAATAGTTTGGCAACGTTGTGTCATTTGCTACAGGCATCGTGTGTCAAGCTCGTCG 474
QY 3017 TTTGGTATGGCTTCATTTCAGCTCCGGTTCCTCAACGATCAAGGCGAGTTACA-TGATCCCC 3075
Db 473 TTTGGTATGGCTTCATTTCAGCTCCGGTTCCTCAACGATCAAGGCGAGTTACATTGATCCCC 414
QY 3076 CATGTTGTCAAAAAAGCGTTAGTCTCTTCGGTCTCCGATCGTTGTGTCAGAAAGTAAGTT 3135
Db 413 CATGTTGTCAAAAAAGCGTTAGTCTCTTCGGTCTCCGATCGTTGTGTCAGAAAGTAAGTT 355
QY 3136 GGGCGCAGTTTATCACTCATGTTATGGCAGCAGTGCATTAATCTTCTTACTCTCATGCC 3195
Db 354 GGGCGCAGTTTATCACTCATGTTATGGCAGCAGTGCATTAATCTTCTTACTCTCATGCC 295
QY 3196 ATCCGTAAGATGCTTTTCTGTGACTGGTGTACTCAACCAAGTCAATCTTGAGAATAGTG 3255
Db 294 ATCCGTAAGATGCTTTTCTGTGACTGGTGTACTCAACCAAGTCAATCTTGAGAATAGTG 235
QY 3256 TATGCGGCGACCGAGTTGCTCTTTGCCGCGCTCAATACGGGATAATACCGCGCCACATAG 3315
Db 234 TATGCGGCGACCGAGTTGCTCTTTGCCGCGCTCAATACGGGATAATACCGCGCCACATAG 175
QY 3316 CAGAAGTTTAAAGTGTCTCATCTTGGAAAAAGTCTTTCGGGCGGAAAACTCTCAAGGAT 3375
```

Db	174	CAGAACTTTAAAGTGCTCATCATTTGGAAAAAGTTCTTCGGGGCGGAAAACTCTCAAGGAT	115
QY	3376	CTTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGC	3435
Db	114	CTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGC	55
QY	3436	ATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAAACAGGAAGCAAAATGCC	3489
Db	54	ATCTTTAACTTTTACCAGCGTTCTGGGTGAGCAAAAAACAGGAAGCAAAATGCC	1
RESULT 4			
AU081137			
LOCUS	AU081137	1067 bp mRNA linear EST 30-JUL-2002	
DEFINITION	AU081137	Oncorhynchus mykiss Kidney infected by infectious hematopoietic necrosis virus	Oncorhynchus mykiss cDNA clone KI2, mRNA sequence.
ACCESSION	AU081137		
VERSION	AU081137.1	GI:6431485	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity	26.1%	Score 955.2; DB 1; Length 1067;	
Matches	982; Conservative	0; Mismatches 8; Indels 2; Gaps 2;	
QY	1494	CGCGTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGTAAGTAATGTTATCCGCTCAAA	1553
Db	78	CGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGTAAGTAATGTTATCCGCTCAAA	137
QY	1554	TTCCACACACATACGACCGGAGCATAAAGTGTAAGACCTGGGTCCTTAATGAGTGA	1613
Db	138	TTCCACACACATACGACCGGAGCATAAAGTGTAAGACCTGGGTCCTTAATGAGTGA	197
QY	1614	GCTAACTCACATTAATTTGGCTTGCCTCACTCCCGCTTTTCCAGTCGGGAAACCTGTGCT	1673
Db	198	GCTAACTCACATTAATTTGGCTTGCCTCACTCCCGCTTTTCCAGTCGGGAAACCTGTGCT	257
QY	1674	GCCAGCTGATTAATGAATTCGCCAAACGCGGGGAGAGCGGTTTGGTATTTGGCGCT	1733
Db	258	GCCAGCTGATTAATGAATTCGCCAAACGCGGGGAGAGCGGTTTGGTATTTGGCGCT	317
QY	1734	CTTCCGCTTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT	1793
Db	318	CTTCCGCTTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT	377
QY	1794	CAGCTCACTCAAAGCGGGTAATACGCTTTATCCACAGAATCAGGGGATACGAGGAAGA	1853
Db	378	CAGCTCACTCAAAGCGGGTAATACGCTTTATCCACAGAATCAGGGGATACGAGGAAGA	437
QY	1854	ACATGTGAGCAAAAAGGCGCAGCAAAAAGGCGCAGAAACCGTAAAAAGGCGGTTGCTGGCGT	1913
Db	438	ACATGTGAGCAAAAAGGCGCAGCAAAAAGGCGCAGAAACCGTAAAAAGGCGGTTGCTGGCGT	497
QY	1914	TTTTTCATAGGCTCGCCCGCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGT	1973
Db	498	VVTTTCATAGGCTCGCCCGCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGT	557
QY	1974	GGCGAAACCCGACAGGACTATAAAGATACCAAGCGCTTTCCCGCTCGAAGCTCCCTCGTGC	2033
Db	558	GGCGAAACCCGACAGGACTATAAAGATACCAAGCGCTTTCCCGCTCGAAGCTCCCTCGTGC	617
QY	2034	GCTCTCTCTTCCGACCCCTGCGCGTTACCGGATACCTGTCCGCCCTTTCTCCCTCGGGAA	2093
Db	618	GCTCTCTCTTCCGACCCCTGCGCGTTACCGGATACCTGTCCGCCCTTTCTCCCTCGGGAA	677
QY	2094	CGTGGCGCTTCTCATAGCTCAGCTGATGATATCTCAGTTTCGGTGTAGGTCTTCGCT	2153
Db	678	GGTGGCGCTTCTCATAGCTCAGCTGATGATATCTCAGTTTCGGTGTAGGTCTTCGCT	737
QY	2154	CAAAGCTGGGCTGTGTGCAAGAAACCCCGCTTTCAGCCGACCGCTTCGCGCTTTATCCGTA	2213
Db	738	CAAAGCTGGGCTGTGTGCAAGAAACCCCGCTTTCAGCCGACCGCTTCGCGCTTTATCCGTA	797
QY	2214	ACTATCGCTTGTAGTCCAAACCGGTAAGACAGACTTATCGGCACTGGCAGCAGCACTG	2273
Db	798	ACTATCGCTTGTAGTCCAAACCGGTAAGACAGACTTATCGGCACTGGCAGCAGCACTG	857
QY	2274	GTAACAGGATTAAGCAGGAGGTATGTAGCGGTGTACAGAGTTCTTGAAGTGTGTGC	2333
Db	858	GTAACAGGATTAAGCAGGAGGTATGTAGCGGTGTACAGAGTTCTTGAAGTGTGTGC	917
QY	2334	CTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGGCTCTGCTGAAGCCAGTTA	2393
Db	918	CTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGGCTCTGCTGAAGCCAGTTA	977
QY	2394	CCTTCGGAAGAGTTGGTAGCTCTTGATCGGCAAAACAAACACCGCTGTGTAGCGGTG	2453
Db	978	CCTTCGGAAGAGAGTTGGTAGCTCTTGATCGGCAAAACAAACCAACGCTGTGTAGCGGTG	1036
QY	2454	GTTTTTTTGTGGCAAGCAGCAGATTACGCGC	2485
Db	1037	TTTTTTTTTTTTC-AGCAGCAGATTACGCGC	1067
RESULT 5			
AJ281480			
LOCUS	AJ281480	1004 bp mRNA linear EST 30-JUN-2000	
DEFINITION	AJ281480	4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P4G8, mRNA sequence.	
ACCESSION	AJ281480		
VERSION	AJ281480.1	GI:6929360	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
COMMENT			

Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
1. .1004
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P4G8"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/clone_lib="Anopheles gambiae immune competent 4A3A"
/note="Vector: p7713D-Pac (Pharmacia) with a modified polylinker; Site 1: EcoRI; Site 2: NotI; sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996) : Normalization and Subtraction: Two approaches To Facilitate Gene Discovery, Genome Research 6, 791-806."

FEATURES
source

ORIGIN

Query Match 25.2%; Score 924; DB 1; Length 1004;
Best Local Similarity 99.0%; Pred. No. 2.9e-260;
Matches 1001; Conservative 1; Mismatches 2; Indels 7; Gaps 7;

1875 AAAAGCCAGGAACCGTAAAGGCGCGTGTGCGCGTTTTCATAGGCTCCGCCCC 1934
DB |||||
1 AAAAGCCAGGAACCGTAAAGGCGCGTGTGCGCGTTTTCATAGGCTCCGCCCC 60
QY |||||
1935 CTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACGAGCTAT 1994
DB |||||
61 CTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACGAGCTAT 120
QY |||||
1995 AAAGATACAGGCGTTTCCCGTGGAGCTCCCTCGTGGCTCTCTGTTCCGACCTGC 2054
DB |||||
121 AAAGATACAGGCGTTTCCCGTGGAGCTCCCTCGTGGCTCTCTGTTCCGACCTGC 180
QY |||||
2055 CGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGGGCTTCTCATAGCT 2114
DB |||||
181 CGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGGGCTTCTCATAGCT 240
QY |||||
2115 CACGCTGTAGGTATCTCAGTTCGGTGTAGTTCGCTCCGCTCAAGCTGGCTGTGTCAG 2174
DB |||||
241 CACGCTGTAGGTATCTCAGTTCGGTGTAGTTCGCTCCGCTCAAGCTGGCTGTGTCAG 300
QY |||||
2175 AACCCCGCTTCAGCCGACCGCTGCGCTTATCCGTTAACTATCGTCTTGAGTCCACCC 2234
DB |||||
301 AACCCCGCTTCAGCCGACCGCTGCGCTTATCCGTTAACTATCGTCTTGAGTCCACCC 360
QY |||||
2235 CGGTAAGACACGACTTATCCCACTGGCAGCAGCCACTGTAAACAGATTAGCAGAGCGA 2294
DB |||||
361 CGGTAAGACACGACTTATCCCACTGGCAGCAGCCACTGTAAACAGATTAGCAGAGCGA 420
QY |||||
2295 GGTATGTAGCGGTGTCTACAGAGTCTTGAAGTGGTGGCTTAACCTACGCTACACTAGAA 2354
DB |||||
421 GGTATGTAGCGGTGTCTACAGAGTCTTGAAGTGGTGGCTTAACCTACGCTACACTAGAA 480
QY |||||
2355 GGACAGTATTTGGTATCTCGGCTCTCTGAGCCAGTACCTTCGGAAAAAGATTGGTA 2414
DB |||||
481 GAACAGTATTTGGTATCTCGGCTCTCTGAGCCAGTACCTTCGGAAAAAGATTGGTA 540
QY |||||
2415 GCTCTTGATCCGCAACCAACACCGCTGGTAGCGGTGTTTTTTTGTTCGACGACG 2474
DB |||||
541 GCTCTTGATCCGCAACCAACACCGCTGGTAGCGGTGTTTTTTTGTTCGACGACG 600
QY |||||
2475 AGATTACGGCGCAAAAAAGGATCTCAAGAAGATCTTTTGATCTTTTCTACGGGCTCG 2534
DB |||||
601 AGATTACGGCGCAAAAAAGGATCTCAAGAAGATCTTTTGATCTTTTCTACGGGCTCG 660
QY |||||
2535 AGCTCAGTGGAAACGAAATCTACGTTTAAAGGATTTTGGTCTATGAGTATCAAAAGGA 2594
DB |||||

Db 661 AGCTCAGTGGAAACGAAACTCAGTTAAGGATTTTGGTCTATGAGTATCAAAAGGA 720
QY 2595 TCTTCCACCTAGATCCTTTTAAATTTAAATGAAGTTTAAATCAATCTAAAGTATATG 2654
Db 721 TCTTCCACCTAGATCCTTTTAAATTTAAATGAAGTTTAAATCAATCTAAAGTATATG 780
QY 2655 AGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGGGACCTATCTCAGCATCT 2714
Db 781 AGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGGGACCA-CTATCTCAGCATCT 839
QY 2715 GTCTATTTCTGTTTCATCATAGTTCCTGCTGCTCCCTCGTGTAGATAACTACCATACGGG 2774
Db 840 GTCTATTTCTGTTTCATCATAGTTCCTGCTGCTCCCTCGTGTAGAT-ACTACCATACGGG 898
QY 2775 AGGGCTTACCATCTGCGCCCGAGTGTGCAATGATACCGGAGACCCACGCTCACCGGCTC 2834
Db 899 AGGGCTTACCATCTGCG-CCAGTGTGCAATGATACCGGAGCA-CCACGCTCA-CGGCTC 955
QY 2835 CAGATTATCAGCAATAAACCGACCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGG 2885
Db 956 CAGATTAT-TCAGCAATAAACCGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1004

RESULT 6
AU081124 1089 bp mRNA linear EST 30-JUL-2002
LOCUS AU081124 Oncorhynchus mykiss Kidney infected by infectious
DEFINITION hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone KG'12,
mRNA sequence.
ACCESSION AU081124 GI:6431472
VERSION AU081124.1
KEYWORDS EST.
SOURCE Oncorhynchus mykiss (rainbow trout)
ORGANISM Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 1089)
AUTHORS Kono, T., Sakai, M. and LaPatra, S.E.
TITLE Expressed Sequence Tag Analysis of Kidney and Gill Tissues from
Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious
Hematopoietic Necrosis Virus
JOURNAL Mar. Biotechnol. 2 (5), 493-498 (2001)
COMMENT Contact: Masahiro Sakai
Faculty of Agriculture
Miyazaki University
1-1 nishi gakukenibandai, Miyazaki, Miyazaki 889-2192, Japan
Email: m.sakai@cc.miyazaki-u.ac.jp.
FEATURES
source
1. .1089
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/clone="KG'12"
/tissue_type="kidney"
/clone_lib="Oncorhynchus mykiss Kidney infected by
infectious hematopoietic necrosis virus"
/note="common name:rainbow trout ; infected by infectious
hematopoietic necrosis virus"

ORIGIN

Query Match 25.2%; Score 923.8; DB 1; Length 1089;
Best Local Similarity 97.2%; Pred. No. 3.3e-260;
Matches 983; Conservative 0; Mismatches 22; Indels 6; Gaps 4;

1494 CCGGCTTGGCGTAATCATGCTCATAGCTGTTTCTGTTGAAATTTTATCCGCTCACAA 1553
QY 79 CGAGCTTGGCGTAATCATGCTCATAGCTGTTTCTGTTGAAATTTTATCCGCTCACAA 138
Db 1554 TTCCACACAACTACGAGCGGAGCATAAAGTGTAAAGCCCTGGGCTGCTTAATGATGA 1613
QY 139 TTCCACACAACTACGAGCGGAGCATAAAGTGTAAAGCCCTGGGCTGCTTAATGATGA 198
Db

QY	1614	GCTAACTCATAATTAATGCGTTGCGCTCACTGCGCGCTTTCCAGTCGGGAACCTGTCTGT	1673
Db	199	GCTAACTCATAATTAATGCGTTGCGCTCACTGCGCGCTTTCCAGTCGGGAACCTGTCTGT	258
QY	1674	GCCAGTGCATTAATGAATTCGCCCAACGCGCGGGAGAGCGGTTGCGTATTGGCGCT	1733
Db	259	GCCAGTGCATTAATGAATTCGCCCAACGCGCGGGAGAGCGGTTGCGTATTGGCGCGCT	318
QY	1734	CTTCCGCTTCTTCCTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGCGAGCGGTAT	1793
Db	319	CTTCCGCTTCTTCCTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGCGAGCGGTAT	378
QY	1794	CAGCTCACTCAAAAGCGGTAATACGGTTATCACAGAATCAGGGGATACGCGAGAAAGA	1853
Db	379	CAGCTCACTCAAAAGCGGTAATACGGTTATCACAGAATCAGGGGATACGCGAGAAAGA	438
QY	1854	ACATGTGAGCAAAAGCCGACGAAAGGCGCAGAAACCGTTAAAGGCGCGGTTGCTGGCGT	1913
Db	439	ACATGTGAGCAAAAGGCGCAGAAAGGCGCAGAAACCGTTAAAGGCGCGGTTGCTGGCGT	498
QY	1914	TTTTCCATAGGCTCCGCGCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGT	1973
Db	499	TTTTCCATAGGCTCCGCGCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGT	558
QY	1974	GGCGAAACCCGACAGGACTATAAAGATACAGGCGGTTTCCCGCTTGGAACTCCCTCGTGC	2033
Db	559	GGCGAAACCCGACAGGACTATAAAGATACAGGCGGTTTCCCGCTTGGAACTCCCTCGTGC	618
QY	2034	GCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTGCGCCCTTCTCCCTTCGGGA	2093
Db	619	GCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTGCGCCCTTCTCCCTTCGGGA	678
QY	2094	GGTGGCGCTTCTCATAGCTCACGCTGTAGTATCTCAGTTTCGCTGTAGTTCGCT	2153
Db	679	GGTGGCGCTTCTCATAGCTCACGCTGTAGTATCTCAGTTTCGCTGTAGTTCGCT	738
QY	2154	CAAAGCTGGCTGTGTGACGAAACCCCGTTTCAGCCGACCGCTTCGCGCTTATCCGGTA	2213
Db	739	CAAAGCTGGCTGTGTGACGAAACCCCGTTTCAGCCGACCGCTTCGCGCTTATCCGGTA	798
QY	2214	ACTATCGTTCGAGTCCAAACCGGTAAGACAGCTTATCGCACCTGGCAGGACCACTG	2273
Db	799	ACTATCGTTCGAGTCCAAACCGGTAAGACAGCTTATCGCACCTGGCAGGACCACTG	858
QY	2274	GTAACAGGATTAGCAGCGAGGTATGTAGGCGGTGCTACAGATTCTTGAAG-TGGTGG	2332
Db	859	GTAACAGGATTAGCAGCGAGGTATGTAGGCGGTGCTACAGATTCTTGAAGTTTGGTGG	918
QY	2333	CCTAACTACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTT	2392
Db	919	CCTAACTACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTT	978
QY	2393	ACCTTCGAAAAAGATTGGTAGCTCTT-GATCCGGGAAACAAACCA---CGCTGGTAG	2448
Db	979	ACCTTCGAAAAAGATTGGTAGCTCTTGGATCCGGGCAACAAACCAACCGCTGGGAAG	1038
QY	2449	CGGTGGTTTTTTTGTTCGAAGCA-GCAGATTACGCGCAGAAAAAAGGAT	2498
Db	1039	GGGTGTCTTTTGTTCGAAGCAGCAGATTACCCGCGAAAAAACCGAT	1089
RESULT 7			
LOCUS	CL021189	1049 bp	DNA linear GSS 31-DEC-2003
DEFINITION	CH216-8A14_RM1.1 CH216 xenopus tropicalis genomic clone CH216-8A14,		
ACCESSION	CL021189		
VERSION	CL021189.1		
KEYWORDS	GSS.		
SOURCE	Xenopus tropicalis (western clawed frog)		
ORGANISM	Xenopus tropicalis		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;		
<hr/>			
Xenopodinae; Xenopus; Silurana.			
1 (bases 1 to 1049)			
Kremitzki C., Carter, J., McPherson, J., Warren, W., Graves, T.,			
Mardis, E. and Wilson, R.			
A physical map of the xenopus tropicalis genome			
Unpublished (2003)			
Contact: Richard K Wilson			
Genome Sequencing Center			
Washington University School of Medicine			
Email: submissions@watson.wustl.edu			
Insert Length: 175000 Std Error: 0.00			
Seq primer: RM1 TACGACTCACTATAGGGAGA			
Class: BAC ends			
High quality sequence start: 43			
High quality sequence stop: 888.			
FEATURES			
source			
1..1049			
/organism="Xenopus tropicalis"			
/mol_type="genomic DNA"			
/strain="Nigerian frog"			
/db_xref="taxon:8364"			
/clone="CH216-8A14"			
/sex="male"			
/cell_line="Stock 248 F7A2, inbred N7"			
/clone_lib="CH216"			
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis			
BAC library"			
ORIGIN			
Query Match 24.7%; Score 902.8; DB 9; Length 1049;			
Best Local Similarity 98.0%; Pred. No. 5.2e-254;			
Matches 935; Conservative 0; Mismatches 17; Indels 2; Gaps 2;			
QY	1486	CGCGCGGCGCGCTGGCGTAATCATGTCTAGCTGTTCTCTGTGAAATGTTATCC	1545
Db	95	CAGGCATCAAGCTTGGCGTAATCATGTCTAGCTGTTCTCTGTGAAATGTTATCC	154
QY	1546	GCTCACAATTCACACAACATACAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTA	1605
Db	155	GCTCACAATTCACACAACATACAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTA	214
QY	1606	ATGAGTAGCTAACTCACTAAATTTGCTTGGCTCACTGCGCGCTTTCCAGTCGGGAAA	1665
Db	215	ATGAGTAGCTAACTCACTAAATTTGCTTGGCTCACTGCGCGCTTTCCAGTCGGGAAA	274
QY	1666	CCTGTCTGCGAGCTGCAATTAATGAATCGCCCAACGCGCGGGAGAGCGGTTTCGGTAT	1725
Db	275	CCTGTCTGCGAGCTGCAATTAATGAATCGCCCAACGCGCGGGAGAGCGGTTTCGGTAT	334
QY	1726	TGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCCGCTCGCGG	1785
Db	335	TGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCCGCTCGCGG	394
QY	1786	AGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATCCAGATATCAGGGGATAACGC	1845
Db	395	AGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATCCAGATATCAGGGGATAACGC	454
QY	1846	AGGAAGACATGTGACAAAGGCCAGCAAAAGCCAGGAACCGTAAAGGCGCGCTT	1905
Db	455	AGGAAGACATGTGACAAAGGCCAGCAAAAGCCAGGAACCGTAAAGGCGCGCTT	514
QY	1906	GCTGGCGTTTTTTCATAGGCTCCGCGCCCTCGAGCATCAAAAAATTCAGCGCTCAAG	1965
Db	515	GCTGGCGTTTTTTCATAGGCTCCGCGCCCTCGAGCATCAAAAAATTCAGCGCTCAAG	574
QY	1966	TCAGAGTGGGAAACCCGACAGGACTATAAGATACAGGGGTTTCCCGCTTGAAGCTC	2025
Db	575	TCAGAGTGGGAAACCCGACAGGACTATAAGATACAGGGGTTTCCCGCTTGAAGCTC	634
QY	2026	CCTGTGCGCTCTCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCC	2085
Db	635	CCTGTGCGCTCTCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCC	694

QY 2086 TTCGGAAGCGTGGCGCTTCTCATAGCTCAGCTGCTAGGTATCTCAGTTCGGTGTAGGT 2145
Db 695 TTCGGAAGCGTGGCGCTTCTCATAGCTCAGCTGCTAGGTATCTCAGTTCGGTGTAGGT 754
QY 2146 CGTTCCCTCAAGCTGGGCTGTGTGACGAACCCCGCTTTCAGCCGACCGCTGCGCCTT 2205
Db 755 CGTTCCCTCAAGCTGGGCTGTGTGACGAACCCCGCTTTCAGCCGACCGCTGCGCCTT 814
QY 2206 ATCCGGTAATATCGTTTGTAGTCCAAACCGGTGAAGACAGCTTATGCCCACTGGCAGC 2265
Db 815 ATCCGGTAATATCGTTTGTAGTCCAAACCGGTGAAGACAGCTTATGCCCACTGGCAGC 874
QY 2266 AGCCACTGTACAGATTAGCAGACGAGGTATGTAGCGGTGCTACAGAGTTCTTTGAA 2325
Db 875 AGCCACTGTAAACAGATTAGCAGACGAGGTATGTAGCGGTGCTACAGAGTTCTTTGAA 934
QY 2326 GTGGTGGCTTAACCTACGGCTACACTAGAAGGAC - AGTATTGTGTATCTGCGCTCTCTGA 2384
Db 935 GTGGTGGCTTAACCTACGGCTACACTAGAAGGACAGTATTGTATTCTGGCTCTCTGA 994
QY 2385 AGCCAGTTACC - TTCGGAAGAAAGAGTTGGTGTAGCTTTGTATCCGGCAAAAC 2437
Db 995 AACCAGTTACCTTCGGAAGAAAGAGTTGGTGTAGCTTTGTATCCGGCAAAAC 1048

RESULT 8
CO487414
LOCUS
DEFINITION Q0227.B7.1.K04 G0022: ROOT XYLEM - mature trees Picea glauca cDNA
clone Q0227.B7_K04 3', mRNA sequence.
ACCESSION CO487414
VERSION
KEYWORDS
SOURCE
ORGANISM Picea glauca (white spruce)

REFERENCE
AUTHORS
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
1 (bases 1 to 928)
Morency, M.-J., Cooke, J., Pavy, N., Parsons, L., Paule, C., Seguin, A.,
Retzel, E., Butterfield, Y., Barber, S., Yang, G., Stott, J.,
Siddiqui, A., Holt, R., Marra, M. and Mackay, J.
Arborea EST sequencing in Picea glauca (white spruce)
Unpublished (2004)
Contact: John Mackay
Centre de Recherche en Biologie Forestiere
Universite Laval
Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4
Fax: 418 656 7493
Email: jmackay@rsvs.ulaval.ca
Center for Computational Genomics and Bioinformatics (CCGB),
University of Minnesota, MN id Identifier: MN5286152 Clone ID:
Q0227.B7.K04 Clones available through: John Mackay, Ph. D.
Professeur adjoint - Assistant professor EMAIL:
jmackay@rsvs.ulaval.ca Centre de Recherche en Biologie Forestiere
(Forest Biology Research Center) Universite Laval Quebec, Quebec
CANADA G1K 7P4
Plate: 7.B7 row: 04 column: K
Seq primer: T7 Primer.
Location/Qualifiers
1..928
/organism="Picea glauca"
/mol_type="mRNA"
/strain="Three trees Unknown"
/db_xref="taxon:3330"
/clone="Q0227.B7.K04"
/sex="Hermaphrodite"
/tissue_type="Differentiating xylem from roots 1 cm in
diameter or larger"
/dev_stage="Non-lignified xylem tissues from larger roots
in early part of growing season (June)."
/lab_host="E. coli DH10B cells"
/clone_lib="GQ022: ROOT XYLEM - mature trees"
/note="Organ: Roots from 9 year old trees measuring

approximately 4.5 m tall, and 10 cm in diameter.: Vector:
pBluescript II SK (+) XR; Site_1: Eco-R1; Site_2: Xho-I;
cDNA was prepared from 5 mg of poly A+ selected RNA and
was directionally ligated into the pBluescript II SK (+)
XR vector (Stratagene), transformed by electroporation
into DH10B cells (in vitro) for propagation"

ORIGIN

Query Match 24.3%; Score 889.8; DB 7; Length 928;
Best Local Similarity 96.0%; Pred. No. 3.4e-250;
Matches 891; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1915 TTTCATAGGCTCCGCGCCCTGACGAGCATCAAAATTCGACGCTCAAGTCAGAGTG 1974
Db 1 TTTCATAGGCTCCGCGCCCTGACGAGCATCAAAATTCGACGCTCAAGTCAGAGTG 60
QY 1975 GCGAAACCCGACAGGACTATAAAGATACAGAGGCTTTCCCTCGAAGCTCCCTCGTCG 2034
Db 61 GCGAAACCCGACAGGACTATAAAGATACAGAGGCTTTCCCTCGAAGCTCCCTCGTCG 120
QY 2035 CTCTCTGTTCGACCCCTGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAG 2094
Db 121 CTCTCTGTTCGACCCCTGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAG 180
QY 2095 CGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTTCGCTC 2154
Db 181 CGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTTCGCTC 240
QY 2155 CAACTGGGCTGTGTGACGAAACCCCTGTCAGCCCGACCGCTCGCTTATCCGGTAA 2214
Db 241 CAACTGGGCTGTGTGACGAAACCCCTGTCAGCCCGACCGCTCGCTTATCCGGTAA 300
QY 2215 CTATCGTCTTGTAGTCCAAACCCGCTAAGACACGACTTATCGCCACTGGCAGCAGCTGG 2274
Db 301 CTATCGTCTTGTAGTCCAAACCCGCTAAGACACGACTTATCGCCACTGGCAGCAGCTGG 360
QY 2275 TAAACAGATTAGCAGACGAGGTATGTAGCGGTGCTACAGATTCTTGAAGTGGTGCC 2334
Db 361 TAAACAGATTAGCAGACGAGGTATGTAGCGGTGCTACAGATTCTTGAAGTGGTGCC 420
QY 2335 TAACTACGGCTACACTAGAGGACAGTATTTGGTATCTCGCTCTCTGTAAGCAGTTAC 2394
Db 421 TAACTACGGCTACACTAGAGGACAGTATTTGGTATCTCGCTCTCTGTAAGCAGTTAC 480
QY 2395 CTTTCGAAAAAGAGTTGGTGTAGCTCTTGTATCCGCAACAAACACCGCTGGTACCGGTGG 2454
Db 481 CTTTCGAAAAAGAGTTGGTGTAGCTCTTGTATCCGCAACAAACACCGCTGGTACCGGTGG 540
QY 2455 TTTTGTGTTCGACGACGAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTT 2514
Db 541 TTTTGTGTTCGACGACGAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTT 600
QY 2515 GATCTTTCTACGGGTCTGACGCTCAGTGGACGAAACTCAGCTTAAGGGATTTGGT 2574
Db 601 GATCTTTCTACGGGTCTGACGCTCAGTGGACGAAACTCAGCTTAAGGGATTTGGT 660
QY 2575 CATGAGATTATCAAAAAAGGATCTTACCTAGATCTTTTAAATTAATAAATTAAGTTTAA 2634
Db 661 CATGAGATTATCAAAAAAGGATCTTACCTAGATCTTTTAAATTAATAAATTAAGTTTAA 720
QY 2635 ATCAATCTAAAGTATATATAGTAAACTTGGTGTGACAGTTACCAATGCTTAATCAGTGA 2694
Db 721 ATCAATCTAAAGTATATATAGTAAACTTGGTGTGACAGTTACCAATGCTTAATCAGTGA 780
QY 2695 GGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCATAGTTGCTGCTCCCGCTCGT 2754
Db 781 GGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCATAGTTGCTGCTCCCGCTCGT 840
QY 2755 GTAGATAACTACGATACGCGAGGCTTACCATCTGCGCCCGAGTGTGCAATGATACCGCG 2814
Db 841 GTNNNTNACTACGATACGCGAGGCTTACCATCTGNNNNNNNGTGTGCAATGATACGCG 900
QY 2815 AGACCCACGCTCACCGGCTCCAGATTAA 2842


```
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp469G045"
/tissue_type="kidney"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="469 (synonym: pkid1)"
/notes="Vector: pSport1_Sfi; Site_1: SfiI; Site_2: SfiI"

ORIGIN
Query Match      23.9%; Score 876.4; DB 7; Length 902;
Best Local Similarity 99.1%; Pred. No. 3e-246; 1; Indels 2; Gaps 2;
Matches 899; Conservative 0; Mismatches 1;

2758 GATAACTACGATACGGAGGGCTTACCACTGCCCCAGTGTGCAATGATACCGGAGA 2817
      |||
901 GATAACTACGATACGGAGGGCTTACCACTGCCCCAGTGTGCAATGATACCGGAGA 842
      |||
2818 CCACGCTCACGGCTCCAGATTTATAGCAATAAACCGACGACCGGAGGCCGAGCG 2877
      |||
841 CCACGCTCACGGCTCCAGATTTTTCAGCAATAAACCGACGACCGGAGGCCGAGCG 782
      |||
2878 CAGAGTGTGCTGCAACTTATCCGCTCCATCCAGTCTATTAAATGTTGCGGGAAGC 2937
      |||
781 CAGAGTGTGCTGCAACTTATCCGCTCCATCCAGTCTATTAAATGTTGCGGGAAGC 722
      |||
2938 TAGAGTAAGTAGTTCGCCAGTTTAATAGTTTGCCAAACGTTGTTGCCATTGCTACAGGCAT 2997
      |||
721 TAGAGTAAGTAGTTCGCCAGTTTAATAGTTTGCG-ACGTTGTTGCCATTGCTACAGGCAT 663
      |||
2998 CTTGTGTCAACGCTCGTCTGTTGGTATGCTTCATTAGTCTCGGTTCCCAAGCATCAAG 3057
      |||
662 CTTGTGTCAACGCTCGTCTGTTGGTATGCTTCATTAGTCTCGGTTCCCAAGCATCAAG 603
      |||
3058 GCGAGTTACATGATCCCCATGTTGTGCAAAAAGCGGTAGTCTCTTGGTCTCTCGAT 3117
      |||
602 GCGAGTTACATGATCCCCATGTTGTGCAAAAAGCGGTAGTCT-CTTGGTCTCTCGAT 544
      |||
3118 CGTTGTGAGAAGTAAGTTGGCGCAGTGTATCACTCATGTTGATGGCAGCACTGCATAA 3177
      |||
543 CGTTGTGAGAAGTAAGTTGGCGCAGTGTATCACTCATGTTGATGGCAGCACTGCATAA 484
      |||
3178 TTCTCTTACTGTCAAGCATCCGTAAGATGCTTTTCTGTGACTGTGAGTACTCAACCAA 3237
      |||
483 TTCTCTTACTGTCAAGCATCCGTAAGATGCTTTTCTGTGACTGTGAGTACTCAACCAA 424
      |||
3238 GTCAATCTGAGATAGTGTATCGGCGACCGAGTGTCTTGTGCGCGCTCAATACGGA 3297
      |||
423 GTCAATCTGAGATAGTGTATCGGCGACCGAGTGTCTTGTGCGCGCTCAATACGGA 364
      |||
3298 TAATACCGCGCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAACGTTCTTCGGG 3357
      |||
363 TAATACCGCGCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAACGTTCTTCGGG 304
      |||
3358 GCGAAACTCTCAAGGATCTTACCGTGTGTAGATCCAGTTCGATGTAAACCACTCGTC 3417
      |||
303 GCGAAACTCTCAAGGATCTTACCGTGTGTAGATCCAGTTCGATGTAAACCACTCGTC 244
      |||
3418 ACCCAACTGATCTTACAGATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAACAG 3477
      |||
243 ACCCAACTGATCTTACAGATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAACAG 184
      |||
3478 AAGGCAAAATCGCGCAAAAGGAATAAGGGGACACGAAATGTTGAATCTCATACT 3537
      |||
183 AAGGCAAAATCGCGCAAAAGGAATAAGGGGACACGAAATGTTGAATCTCATACT 124
      |||
3538 CTTCTCTTTTCAATATTATTGAGCATTTATCAGGGTTATTTGCTCATGAGCGGATACAT 3597
      |||
123 CTTCTCTTTTCAATATTATTGAGCATTTATCAGGGTTATTTGCTCATGAGCGGATACAT 64
      |||
3598 ATTGTAATGTAATTAAGAAATAAACAATAGGGGTTCCGCGCACATTTCCCGGAAAGT 3657
      |||
```

63 ATTGAATGTTATTAGAAAAATAAACAAATAGGGTTCCGCGCACATTTCCCGGAAAGT 4

3658 GC 3659

3 GC 2

RESULT 11

LOCUS BG838279/c

DEFINITION Glycine claudestina 935 bp mRNA linear EST 25-MAY-2001

ACCESSION BG838279

VERSION BG838279.1 GI:14204601

KEYWORDS EST.

SOURCE Glycine claudestina

ORGANISM Glycine claudestina

REFERENCE 1 (bases 1 to 935)

AUTHORS Singh, J.A., Farah, S., Chapados, J., Couroux, P., De Moors, A., Harris, L.J., Hattori, J.I., Ouellet, T., Robert, L.S., Spratt, D. and Tinker, N.A.

TITLE Expressed Sequence Tags from Cold-Stressed Glycine claudestina Seedlings

JOURNAL Unpublished (2001)

COMMENT Contact: Singh, J.A.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A 0C6, Canada
Tel: (613) 759-1662
Fax: (613) 759-1701
Email: singhja@agr.gc.ca.

FEATURES

source

1..935

/organism="Glycine claudestina"

/mol_type="mRNA"

/cultivar="1035"

/db_xref="taxon:45687"

/clone="Gc01_10e07"

/tissue_type="Leaves, stem"

/clone_lib="Gc01_AAF_Coldest_stressed_Glycine_claudestina"

/note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: EcoRI; Site 2: XhoI; Plants incubated at 2 degrees under 12 hours of light/day. Harvested after only 2-3 days of cold treatment. cDNA was prepared with the Uni-Zap cDNA kit from Stratagene. Eco RI adapters were linked followed by digest with Xho I/Eco RI and ligated to pBluescript."

ORIGIN

Query Match 23.8%; Score 870.2; DB 4; Length 935;

Best Local Similarity 98.1%; Pred. No. 2.1e-244;

Matches 874; Conservative 14; Mismatches 2; Indels 1; Gaps 1;

2773 GGAGGGCTTACCATCTGGCCCGAGTCTGCAATGATACCGGAGCCAGCTCACCGC 2832

928 GGGGGCTTACCATCTGGCCCGAGTCTGCAATGATACCGGAGCCAGCTCACCGC 869

2833 TCCAGATTTTATCAGCAATAAACCGACGACCGGAGGCCGAGAGTGGTCTGTC 2892

868 YCCAGATTTTATCAGCAATAAACCGACGACCGGAGGCCGAGAGTGGTCTGTC 809

2893 AACTTTATCCGCTCCATCCAGTCTATTATTTGTTCCGGGAGCTAGAGTAGTTC 2952

808 AACTTTATCCGCTCCATCCAGTCTATTATTTGTTCCGGGAGCTAGAGTAGTTC 749

2953 GCGAGTTATAGTTTGGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGCAGCTC 3012

748 GCGAGTTATAGTTTGGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGCAGCTC 689

```
QY 3013 GTCTGTTGGTATGGCTTCATTTCAGCTCCGGTTCCCAACGATCAAGCGAGTTACATGATC 3072
DB |||||
DB 688 GTCTGTTGGTATGGCTTCATTTCAGCTCCGGTTCCCAACGATCARGCGAGTTACATGATC 629
QY 3073 CCCATGTTGTGCAAAAAGCGGTTAGCTTCCTTCGGTCTCCGATCGTTGTGTCAGAAATA 3132
DB |||||
DB 628 CCCATGTTGTGCAAAAAGCGGTTAGCTTCCTTCGGTCTCCGATCGTTGTGTCAGAAATA 569
QY 3133 GTT-GGCCGACGTGTTATCACTCATGTTATGGCAGCATGCTAATTTCTTACTGTCA 3191
DB |||||
DB 568 GTTGGGCGCAGTGTATCACTCATGTTATGGCAGCATGCTAATTTCTTACTGTCA 509
QY 3192 TCCCATCCGTAAGATGCTTTTCTGTGACGTGGTGAGTACTCAACCAAGTCATTCTCAGAAAT 3251
DB |||||
DB 508 TCCCATCCGTAAGATGCTTTTCTGTGACGTGGTGAGTACTCAACCAAGTCATTCTCAGAAAT 449
QY 3252 AGTGTATCGGCGACCGAGTTGCTCTTTCGCCGGCTCAATACGGGATAATACCGGCCAC 3311
DB |||||
DB 448 AGTGTATCGGCGACCGAGTTGCTCTTTCGCCGGCTCAATACGGGATAATACCGGCCAC 389
QY 3312 ATAGCAGAACTTTAAAGTGCTCATCATTTGGAAGAGTTCTTCGGGGGAAACCTCTCA 3371
DB |||||
DB 388 ATAGCAGAACTTTAAAGTGCTCATCATTTGGAAGAGTTCTTCGGGGGAAACCTCTCA 329
QY 3372 GGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTCGTGACCCCACTGATCTT 3431
DB |||||
DB 328 GGATCTTACCGCTGTTGAGATCCAGTTCGATGTAMCCCACTCGTGACCCCACTGATCTY 269
QY 3432 CAGCATCTTTTACTTTACCAAGCGTTTCGGGTGAGCAAAAACAGAGAAAGCGCAATCGC 3491
DB |||||
DB 268 CAGCATCTTTTACTTTACCAAGCGTTTCGGGTGAGCAAAAACAGAGAAAGCGCAATCGC 209
QY 3492 CAAAAAGGATATAGGCGGACACGGAATGTGAATCTACTACTCTCTCTCTTTTCAAT 3551
DB |||||
DB 208 CAAAAAGGATATAGGCGGACACGGAATGTGAATCTACTACTCTCTCTCTTTTCAAT 149
QY 3552 ATTATTGAAGCATTTATCAGGGTATTGTCTCATGAGCGGATACATATTGAATGTATT 3611
DB |||||
DB 148 ATTATTGAAGCATTTATCAGGGTATTGTCTCATGAGCGGATACATATTGAATGTATT 89
QY 3612 AGAAAAATAAACAAATAGGGTTCGCGCACATTTCCCGGAAAGTGCCAC 3662
DB |||||
DB 88 AGAAAAATAAACAAATAGGGTTCGCGCACATTTCCCGGAAAGTGCCAC 38

RESULT 12
CF269652
LOCUS
DEFINITION
Fvylcold844 Fragilariopsis cylindrus SMART cDNA library (Clontech)
Fragilariopsis cylindrus cDNA clone Antarctic 5', mRNA sequence.
ACCESSION
CF269652
VERSION
CF269652.1 GI:33631539
SOURCE
Fragilariopsis cylindrus
ORGANISM
Fragilariopsis cylindrus
Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
Bacillariophycidae; Bacillariales; Bacillariaceae; Fragilariopsis.
1 (bases 1 to 1073)
Mock,T. and Valentin,K.
EST analysis of freezing tolerance in the Antarctic diatom
Fragilariopsis cylindrus: Detection of numerous cold adaptation
related genes and gene transfer events
Unpublished (2003)
Contact: Mock T
Biological Oceanography
Alfred-Wegener-Institute for Polar and Marine Research
Am Handelshafen 12, D-27570 Bremerhaven, Germany
Tel.: +49 471 4831 1893
Fax: +49 471 4831 1425
Email: tmock@awi-bremerhaven.de
sequence with unknown function
PCR PRIMERS
```

```
FORWARD: 5' lambdatriplex2
BACKWARD: 3' lambdatriplex2
Seq primer: ctcggaagcgcgcattgtgtgtgt.
FEATURES
location/Qualifiers
source
1..1073
/organism="Fragilariopsis cylindrus"
/mol_type="mRNA"
/strain="Antarctic"
/db_xref="taxon:186039"
/clone="Antarctic"
/clone_lib="Fragilariopsis cylindrus SMART cDNA library
(Clontech)"
/note="Vector: pTriplex2; total polyA was used for
first-strand synthesis with SMART IV oligos and CDS
III/3'PCR primer. Double strand cDNA synthesis was done by
LD PCR using the following program: 95oC for 5 min
denaturation and subsequent 20 cycles at 95oC (2min) and
68oC (6min). After SfiI digestion the cDNA was
fractionated with CHROMA Spin-400 columns. These cDNAs
were ligated overnight into pTriplex2 vectors."
```

ORIGIN

```
Query Match 23.6%; Score 865; DB 7; Length 1073;
Best Local Similarity 96.2%; Pred. No. 7.3e-243;
Matches 927; Conservative 0; Mismatches 30; Indels 7; Gaps 4;
QY 1631 GGGTTGCCCTCAGTCCCGCTTCCAGTCGGGAACCTGTCGTCAGCTGCATTAATGA 1690
DB |||||
DB 1 GCTTCTTGCTACTGCCCGCTTTCAGTCGGGAACCTGTCGTCAGCTGCATTAATGA 60
QY 1691 ATCGSCCAACCGCGGGGAGAGCGGTTTGGGTATTGGCGCTCTTCCGCTTCCTCGCTC 1750
DB |||||
DB 61 ATCGSCCAACCGCGGGGAGAGCGGTTTGGGTATTGGCGCTCTTCCGCTTCCTCGCTC 120
QY 1751 ACTGACTCGCTCGCTCGCTGCTGCGCTGCGGAGCGGTATCAGCTCACTCAAGGCG 1810
DB |||||
DB 121 ACTGACTCGCTCGCTCGCTCTTTCCGCTGCGGAGCGGTATCAGCTCACTCAAGGCG 180
QY 1811 GTAATACGGTTATCCACAGAAATCAGGGATTAACGAGGAAGAAACATGTGACAAAGGC 1870
DB |||||
DB 181 GTAATACGGTTATCCACAGAAATCAGGGATTAACGAGGAAGAAACATGTGACAAAGGC 240
QY 1871 CAGCAAAAGGCGAGGAACCGTAAAGGCGCGTTTGGCTGGCGTTTTTTCATAGGCTCGCG 1930
DB |||||
DB 241 CAGCAAAAGGCGAGGAACCGTAAAGGCGCGTTTGGCTGGCGTTTTTTCATAGGCTCGCG 300
QY 1931 CCCCTGACGAGCATCAAAAATCGACCTCAAGTCAGAGTGGCGAAACCCGACAGA 1990
DB |||||
DB 301 CCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGA 360
QY 1991 CTATAAGATACCGAGCGGTTTCCCGCTCGAAGCTCCCTCGTCGCTCTCTCTTCCGACC 2050
DB |||||
DB 361 CTATAAGATACCGAGCGGTTTCCCGCTCGAAGCTCCCTCGTCGCTCTCTCTTCCGACC 420
QY 2051 CTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTCGCGTTTCTCAT 2110
DB |||||
DB 421 CTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTCGCGTTTCTCAT 480
QY 2111 AGCTCACGCTGTAGGTATCTCAGTTTCGCTGTAGTTCGTTCCGTCGAAGTGGGCTGTGTG 2170
DB |||||
DB 481 AGCTCACGCTGTAGGTATCTCAGTTTCGCTGTAGTTCGTTCCGTCGAAGTGGGCTGTGTG 540
QY 2171 CACGAACCCCGTTTCAGCCGACCGCTTATCCGGTAACTATCTGCTTCAGTGCC 2230
DB |||||
DB 541 CACGAACCCCGTTTCAGCCGACCGCTTATCCGGTAACTATCTGCTTCAGTGCC 600
QY 2231 AACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGA 2290
DB |||||
DB 601 AACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGA 660
QY 2291 GCGAGGTATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCT--AATACGCGCTACA 2348
DB |||||
DB 661 GCGAGGTATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACNGGCTACA 720
```

```
QY 2349 CTAGAAGGACA-GTATTGTGTATCT-GCGCTCTGTGTAAGCC---AGTTACCTTCGGAAA 2403
|||
Db 721 CTAGAAGAACANGATATTGTGTATCTNGCGCTCTGTCTNGAAGCCAGNNTTACCNNTTCGAAA 780
|||
QY 2404 AAGAGTTGGTAGCTCTTGATCCCGCAAAACAAACACCGCTGGTAGCGGTGTTTTTGT 2463
|||
Db 781 AAGAGTTGGTAGCTCTTGATCCCGCAAAACAAACACCGCTGGTAGCGGTGTTTTTGT 840
|||
QY 2464 TTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAAGATCCTTTGATCTTTTC 2523
|||
Db 841 TTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAAGATCCTTTGATCTTTTC 900
|||
QY 2524 TACGGGGTCTGACGCTCAGTGAACGAAACTCACTCAAGGATTTTGGTCAAGAGATT 2583
|||
Db 901 TACGGGGTCTGACGCTCAGTGAACGAAACTCACTCAAGGATTTTGGGATGAGACA 960
|||
QY 2584 ATCA 2587
|||
Db 961 ATAA 964
|||

RESULT 13
LOCUS CL076016
DEFINITION CH216-138F20_RM1.1 CH216 Xenopus tropicalis genomic clone
ACCESSION CH216-138F20, genomic survey sequence.
VERSION CL076016
KEYWORDS CL076016.1 GI:40531929
SOURCE GSS.
ORGANISM Xenopus tropicalis (western clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 885)
AUTHORS Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: RM1 TACGACTCATATAGGAGA
Class: BAC ends
High quality sequence start: 11
High quality sequence stop: 810.

FEATURES
source
1..885
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-138F20"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/notes="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

ORIGIN
Query Match 23.4%; Score 857.6; DB 9; Length 885;
Blast Local Similarity 99.3%; Pred. No. 1e-240; 5; Indels 1; Gaps 1;
Matches 871; Conservative 0; Mismatches 5;

QY 1535 AATTGTATCCGCTCAATTCACACATACGAGCGGAGCATATAAGTGTAAAGCC 1594
|||
Db 5 AATTGTATCCGCTCAATTCACACATACGAGCGGAGCATATAAGTGTAAAGCC 64
|||
QY 1595 TGGGTGCTTATGAGTGAAGTCACTCAATTAATTGGTGGCTCACTGCCCGCTTTC 1654
|||
```

```
Db 65 TGGGTGCTTATGAGTGAAGTCACTCAATTAATTGGTGGCTCACTGCCCGCTTTC 124
|||
QY 1655 CAGTCCGGAACCTGTCTGTCAGCTGCATTAATGAATCGGCAACCGCGCGGAGAGGC 1714
|||
Db 125 CAGTCCGGAACCTGTCTGTCAGCTGCATTAATGAATCGGCAACCGCGCGGAGAGGC 184
|||
QY 1715 GGTTCGCTATTGGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTCGGCTCGGCTT 1774
|||
Db 185 GGTTCGCTATTGGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTCGGCTCGGCTT 244
|||
QY 1775 CGGCTCGGCGAGCGGTATCAGCTCACTCAAGCGGCTTAATCGGTTATCCACAGAATCA 1834
|||
Db 245 CGGCTCGGCGAGCGGTATCAGCTCACTCAAGCGGCTTAATCGGTTATCCACAGAATCA 304
|||
QY 1835 GGGGATAACCGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGCCAGGAAACCGTAAA 1894
|||
Db 305 GGGGATAACCGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGCCAGGAAACCGTAAA 364
|||
QY 1895 AAGGCCGCTTGTCTGGCGCTTTTTCATAGGCTCCGCCCCCTGACGAGCATCAAAAT 1954
|||
Db 365 AAGGCCGCTTGTCTGGCGCTTTTTCATAGGCTCCGCCCCCTGACGAGCATCAAAAT 424
|||
QY 1955 CGACGCTCAAGTCAGAGTGGGAAACCGCAGAGGACTATAAGATACACAGGCGTTTCCC 2014
|||
Db 425 CGACGCTCAAGTCAGAGTGGGAAACCGCAGAGGACTATAAGATACACAGGCGTTTCCC 484
|||
QY 2015 CTTGGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCTCCGCTTACCGGATACCTGTCC 2074
|||
Db 485 CTTGGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCTCCGCTTACCGGATACCTGTCC 544
|||
QY 2075 GCTTTCTCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAGCTGCTAGGTATCTCAGT 2134
|||
Db 545 GCTTTCTCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAGCTGCTAGGTATCTCAGT 604
|||
QY 2135 TGGGTGCTAGTCTCGTCCGCTCAAGCTGGGCTGTGTCAGCAACCCCGCTTACGCGCGAC 2194
|||
Db 605 TGGGTGCTAGTCTCGTCCGCTCAAGCTGGGCTGTGTCAGCAACCCCGCTTACGCGCGAC 664
|||
QY 2195 CGCTGCGCTTATCCGCTTAAGTATCTGTTGAGTCCAAACCCCGGTAAGACACGCTTATCG 2254
|||
Db 665 CGCTGCGCTTATCCGCTTAAGTATCTGTTGAGTCCAAACCCCGGTAAGACACGCTTATCG 724
|||
QY 2255 CCAGTGGCAGCAGCCACTGCTTAACAGGATAGCAGAGCAGGTATGTAGCGGTGCTACA 2314
|||
Db 725 CCAGTGGCAGCAGCCACTGCTTAACAGGATAGCAGAGCAGGTATGTAGCGGTGCTACA 784
|||
QY 2315 GAGTTCTTGAAGTGGTGGCTAACTACGCTACACTAGAGGACAGTATTTGTTATCTGC 2374
|||
Db 785 GAGTTCTTGAAGTGGTGGCTAACTACGCTACACTAGAGGACAGTATTTGTTATCTGC 844
|||
QY 2375 GCTCTGCTGAAGCCAGTTTACCTTCGGAAAAAGAGTTG 2411
|||
Db 845 GCTCTGCTGAAG-CAGTTACCTTCGGAAAAAGAGTTG 880
|||

RESULT 14
LOCUS AL042026/c
DEFINITION DKF2p434E111.r1.434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION DKF2p434E111.5, mRNA sequence.
VERSION AL042026
KEYWORDS AL042026.1 GI:5421372
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 841)
AUTHORS Poustka, A., Klein, M., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
JOURNAL EST (Poustka, et al.)
COMMENT Unpublished (1999)
MIPS Contact: MIPS
```


Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No si sequence available.
This clone (DKFZq434E11) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de.

FEATURES
source

```

1. .841
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp434E111"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/notes="Vector: pSPort1; Site 1: NotI; Site 2: SalI"

```

ORIGIN

Query Match	23.0%	Score	841;	DB	1;	Length	841;		
Best Local Similarity	100.0%	Pred. No.	8.1e-236;						
Matches	841;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	2788	TGCGCCAGTCTGCAATGATACCGCGAGACCCACACGCTCACCGCTCCAGATTATCAGC	2847						
DB	841	TGCGCCAGTCTGCAATGATACCGCGAGACCCACACGCTCACCGCTCCAGATTATCAGC	782						
QY	2848	AATAAACCAGCCAGCCGGAAGGGCCGAGCCGAGAAGTGGTCTTGCAACTTTATCCGGCTC	2907						
DB	781	AATAAACCAGCCAGCCGGAAGGGCCGAGCCGAGAAGTGGTCTTGCAACTTTATCCGGCTC	722						
QY	2908	CATCCAGTCTATTAAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTGCGCAGTTAATAGTTT	2967						
DB	721	CATCCAGTCTATTAAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTGCGCAGTTAATAGTTT	662						
QY	2968	GCGCAAGTGTGTTGCCATTGCTACAGCATCGTGGTGTACGGTCTGTCGTTGGTATGGC	3027						
DB	661	GCGCAAGTGTGTTGCCATTGCTACAGCATCGTGGTGTACGGTCTGTCGTTGGTATGGC	602						
QY	3028	TTCAATTCAGCTCCGGTTCCTCAACCGATCAAGCGCAGTTTACATGATCCCCCATGTTGTC	3087						
DB	601	TTCAATTCAGCTCCGGTTCCTCAACCGATCAAGCGCAGTTTACATGATCCCCCATGTTGTC	542						
QY	3088	AAAAGCGGTTAGTCTCTCGGTCTCCGATCGTTGTGTCAGAAAGTGGCCCGCAGTGT	3147						
DB	541	AAAAGCGGTTAGTCTCTCTCGGTCTCCGATCGTTGTGTCAGAAAGTGGCCCGCAGTGT	482						
QY	3148	ATCACTCATGGTTATGGCAGCAGCTGCATTAATCTCTTACTGTGTCATGCCATCCGTAA	3207						
DB	481	ATCACTCATGGTTATGGCAGCAGCTGCATTAATCTCTTACTGTGTCATGCCATCCGTAA	422						
QY	3208	CTTTTCTGTGACTGGTGAGTACTCAACCAAGTCAATCTGAGAATAGTGTATCGGGCAGC	3267						
DB	421	CTTTTCTGTGACTGGTGAGTACTCAACCAAGTCAATCTGAGAATAGTGTATCGGGCAGC	362						
QY	3268	GAGTTGCTCTTGCCCGCGGTCAATA CGGGATAATACCGCGCCACATAGCAGAACTTTAA	3327						
DB	361	GAGTTGCTCTTGCCCGCGGTCAATA CGGGATAATACCGCGCCACATAGCAGAACTTTAA	302						
QY	3328	AGTGCTCATCTTGGAAAAAGCTTCTCGGGCGGAAAACTCTCAAGGATCTTACCGCTGTT	3387						
DB	301	AGTGCTCATCTTGGAAAAAGCTTCTCGGGCGGAAAACTCTCAAGGATCTTACCGCTGTT	242						
QY	3388	GAGATCCAGTTGCATGTACCCACTCGTGCACCCAACTGATCTTTCAGCATCTTTTACTTT	3447						
DB	241	GAGATCCAGTTGCATGTACCCACTCGTGCACCCAACTGATCTTTCAGCATCTTTTACTTT	182						
QY	3448	CACGACGGTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAAATCGCGCAAAAAAGGAATAAG	3507						

Db	181	CACCAGCGTTTCTGGGGTACGCAAAAAACAGGAAGCGCAAAATGCCGCAAAAAAGGAATAAG	122
Qy	3508	GGCGACACGGAATGTTGAATACTCATACACTCTCTTTTCAATATATTATTGAAGCATTTTA	3567
Db	121	GGCGACACGGAATGTTGAATACTCATACACTCTCTTTTCAATATATTATTGAAGCATTTTA	62
Qy	3568	TCAGGGTTATTGTCTCATGACGCGATACATATTTGAATGTATTTAGAAAAATAAACAAAT	3627
Db	61	TCAGGGTTATTGTCTCATGACGCGATACATATTTGAATGTATTTAGAAAAATAAACAAAT	2
Qy	3628	A 3628	
Db	1	A 1	
RESULT 15			
BM438950/c			
LOCUS	BM438950	854 bp	linear EST 31-JAN-2002
DEFINITION	IPdVr00491 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA		
ACCESSION	BM438950		
VERSION	BM438950.1	GI:18460672	
KEYWORDS	EST.		
SOURCE	Ictalurus punctatus (channel catfish)		
ORGANISM	Ictalurus punctatus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;		
TITLE	Ictaluridae; Ictalurus.		
JOURNAL	1 (bases 1 to 854)		
COMMENT	Feng,J., Kukutkas,H., Kocabas,A., Li,P. and Liu,Z. Transcriptome of channel catfish (Ictalurus punctatus): initial analysis of expressed sequence tags from the liver Unpublished (2002) Contact: Liu ZJ The Fish Molecular Genetics and Biotechnology Laboratory, Department of Fisheries and Allied Aquacultures and Program of Cell and Molecular Biosciences Auburn University 203 Swingle Hall, Auburn University, Auburn, AL 36849, USA Tel: 334 844 4054 Fax: 334 844 9208 Email: zliu@acesag.auburn.edu Seq primer: M13 Reverse. Location/Qualifiers 1. .854 /organism="Ictalurus punctatus" /mol_type="mRNA" /db_xref="taxon:7998" /clone_lib="Liver cDNA library" /notes="Organ: Liver; Vector: pSport1; Site_1: NotI; Site_2: SalI"		
FEATURES			
source			
ORIGIN			
Query Match	22.5%;	Score 824.6;	DB 4; Length 854;
Best Local Similarity	99.3%;	Pred. No. 5.6e-231;	
Matches 849;	Conservative 0;	Mismatches 4;	Indels 2; Gaps 2;
Qy	2636	TCAATCTAAAGTATATATAGAGTAAACTTGGCTGACAGTTACCAATGCTTAAATCAGTGAG	2695
Db	854	TCAATCTCAAGCATATATAGTAAACTTGGCTCTGACAGTTACCAATGCTTAAATCAGTGAG	795
Qy	2696	GCACCTATCTCAGCGATCTGCTATTTTCGTTTCATCCATAGTTGCCGTCGATCTCCCGTCGTCG	2755
Db	794	GCACCTATCTCAGCGATCTGCTATTCGTTTCATCCATAGTTGCCGTCGATCTCCCGTCGTCG	735
Qy	2756	TAGATAACTACGATACGGGAGGGCTTACCATCTGGGCCCAAGTGTCTCAATGATACCGCGA	2815
Db	734	TAGATAACTACGATACGGGAGGGCTTACCATCTGGGCCCAAGTGTCTCAATGATACCGCGA	675
Qy	2816	GACCCACGCTCACCGCTCCAGATTTATCAGCAATAAACAGCAGCGCGAAGGGCCGAG	2875
Db	674	GACCCACGCTCACCGCTCCAGATTTATCAGCAATAAACAGCAGCGCGAAGGGCCGAG	615

2976	CGCAGAAAGTGGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTCCGGGAA	29335
614	CGCAGAAAGTGGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTCCGGGAA	555
2936	GCTAGAGTAAGTAGTTCCGCAAGTTAAATAGTTTGGCGCAACGTTGTTGCCATTGCTACAGGC	2995
554	GCTAGAGTAAGTAGTTCCGCAAGTTAAATAGTTTGGCGCAACGTTGTTGCCATTGCTACAGGC	495
2996	ATCGTGGTGTCACGCTCGTCGTTGGTATAGGCTTTCAATTCAGTCCCGTTCCCAACGATCA	3055
494	ATCGTGGTGTCACGCTCGTCGTTGGTATAGGCTTTCAATTCAGTCCCGTTCCCAACGATCA	435
3056	AGCGGAGTTACATGATCCCCATGTTGTCGAAAAAAGCGGTTAGCTCCTTCCGGTCCCTCG	3115
434	AGCGGAGTTACATGATCCCCATGTTGTCGAAAAAAGCGGTTAGCTCCTTCCGGTCCCTCG	376
3116	ATCGTTCTCAGAAAGTAAGTTGGCCGCAAGTGTATCACTCATGGTTATGGCAGCACTGCAT	3175
375	ATCGTTCTCAGAAAGTAAGTTGGCCGCAAGTGTATCACTCATGGTTATGGCAGCACTGCAT	316
3176	AATTCTCTTATGTGCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTACTCAACC	3235
315	AATTCTCTTATGTGCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTACTCAACC	256
3236	AAGTCATCTTCGAAATAGTGTATGCGCGCACCGAGTTGCTCTTGCCCGCGCTCAATACGG	3295
255	AAGTCATCTTCGAAATAGTGTATGCGCGCACCGAGTTGCTCTTGCCCGCGCTCAATACGG	196
3296	GATAATACCGCGGCACATAGCAGAACTTTAAAAGTGTCTCATTTGGAAAACGTTCTTCG	3355
195	GATAATACCGCGGCACATAGCAGAACTTTAAAAGTGTCTCATTTGGAAAACGTTCTTCG	136
3356	GGCGGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTAA-CCCACTCG	3414
135	GGCGGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTAAAGCCCACTCG	76
3415	TGCACCAACCTGATCTTTCAGCATCTTTTACTTTTCACCACGCGTTTCTGGGTGAGCAAAAC	3474
75	TGCACCAACCTGATCTTTCAGCATCTTTTACTTTTCACCACGCGTATCTGGGTGAGCAAAAC	16
3475	AGGAAGCAAAATGCG	3489
15	AGGAAGCAAAATGCG	1

Search completed: June 5, 2005, 22:55:16
Job time : 10715.2 secs

This Page Blank (uspto)